

**FIGURE 1**

GTTACTCGGTGGTGGCGGAGTCTACGGAAGCCGTTTTTCGCTTCACTTTTCCTGGCTGTAGAGC  
GCTTTCCCCCTGGCGGGTGAGAGTGCAGAGACGAAGGTGCGAGATGAGCACTATGTTGCGCGGA  
CACTCTCCTCATCGTTTTTTATCTCTGTGTGCACGGCTCTGCTCGCAGAGGGCATAACCTGGGT  
CCTGGTTTACAGGACAGACAAGTACAAGAGACTGAAGGCAGAAGTGGAACACAGAGTAAAAA  
ATTGGAAGAAGAAGGAAACAATAACAGAGTCAGCTGGTCGACAACAGAAAAAGAAAATAGA  
GAGACAAGAAGAGAACTGAAGAATAACAACAGAGATCTATCAATGGTTCGAATGAAATCCAT  
GTTTGCTATTGGCTTTTGTCTTACTGCCCTAATGGGAATGTTCAATTCATATTTGATGGTAG  
AGTGGTGGCAAAGCTTCCTTTTACCCCTCTTCTTACATCCAAGGACTGTCTCATCGAAATCT  
GCTGGGAGATGACACCACAGACTGTTTCCTTCATTTTCTGTATATTCTCTGTACTATGTCGAT  
TCGACAGAACATTCAGAAGATTCTCGGCCTTGCCCTTCACGAGCCGCCACCAAGCAGGCAGG  
TGGATTTCTTGCCCCACCACCTCCTTCTGGGAAGTTCTCTTGAACTCAAGAACTCTTTATTTT  
CTATCATTCCTTCTAGACACACACATCAGACTGGCAACTGTTTTGTAGCAAGAGCCATAGG  
TAGCCTTACTACTTGGGCCTCTTCTAGTTTTTGAATTATTTCTAAGCCTTTTGGGTATGATTA  
GAGTGAAAATGGCAGCCAGCAAACCTTGATAGTGCTTTTGGTCCTAGATGATTTTTATCAAATA  
AGTGGATTGATTAGTTAAGTTCAGGTAATGTTTATGTAATGAAAAACAAATAGCATCCTTCTT  
GTTTCATTTACATAAGTATTTTCTGTGGGACCGACTCTCAAGGCACTGTGTATGCCCTGCAAG  
TTGGCTGTCTATGAGCATTTAGAGATTTAGAAGAAAAATTTAGTTTGTTTAACCTTGTAAC  
GTTTGTTTTGTTGTTGTTTTTTTTTCAAGCCAAATACATGACATAAGATCAATAAAGAGGCCA  
AATTTTGTAGCTGTTTTATGTACAAGGAGAGATCTGTTTCATTTTGTGTTTGCCGTATTTCTAGA  
TATAAGTTTGTAGCATGGGCCAGGAAGGACTAAAATAAAAGTTTTTAAGGTACAAAAAAAAAAAA  
AAAA

## **FIGURE 2**

MSTMFADTLLIVFISVCTALLAEGITWVLVYRTDKYKRLKAEVEKQSKKLEKKKETITESAGR  
QOKKKIERQEEKLKNNNRDLSMVRMKSMFAIGFCFTALMGMFNSIFDGRVVAKLPFTPLSYIQ  
GLSHRNLLGDDTTDCSFIFLYILCTMSIRQNIQKILGLAPSRAATKQAGGFLGPPPPSGKFS

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **N-myristoylation sites.**

amino acids 103-109, 163-169

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 53-57

bioRxiv preprint doi: <https://doi.org/10.1101/000001>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 3**

AGCCGGGGGCGGGTTTGAAGACGCGTCGTTGGGTTTTGGAGGCCGTGAAACAGCCGTTTGAGT  
TTGGCTGCGGGTGGAGAACGTTTGTGAGGGGCCCGCCAAGAAGGAGGCCCGCCTGTTACG**AT**  
**G**GTGTCCATGAGTTTCAAGCGGAACCGCAGTGACCGGTTCTACAGCACCCGGTGCTGCGGCTG  
TTGCCATGTCCGCACCGGGACGATCATCCTGGGGACCTGGTACATGGTAGTAAACCTATTGAT  
GGCAATTTTGCTGACTGTGGAAGTGACTCATCCAACTCCATGCCAGCTGTCAACATTGAGTA  
TGAAGTCATCGGTAATTACTATTCTGCTGAGAGAATGGCTGATAATGCCTGTGTTCTTTTTGC  
CGTCTCTGTTCTTATGTTTATAATCAGTTCAATGCTGGTTTTATGGAGCAATTTCTTATCAAGT  
GGGTTGGCTGATTCCATTCTTCTGTTACCGACTTTTTGACTTCGTCCTCAGTTGCCTGGTTGC  
TATTAGTTCTCTCACCTATTTGCCAAGAATCAAAGAATATCTGGATCAACTACCTGATTTTCC  
CTACAAAGATGACCTCCTGGCCTTGGACTCCAGCTGCCTCCTGTTCAATTGTTCTTGTGTTCTT  
TGCTTATTTCATCATTTTTTAAGGCTTATCTAATTAAGTGTGTTTGGAACTGCTATAAATACAT  
CAACAACCGAAACGTGCCGGAGATTGCTGTGTACCCTGCCTTTGAAAGCACCTCCTCAGTACG  
TTTTGCCAACCTATGAAATGGCCGTGAAAATGCCTGAAAAAGAACCACCACCTCCTTACTTAC  
CTGCCTGAAGAAATTCTGCCTTTGACAATAAATCCTATACCAGCTTTTTGTTTGTTTATGTTA  
CAGAATGCTGCAATTCAGGGCTCTTCAAACCTGTTTGATATAAAATATGTTGTCTTTTGTTTA  
AGCATTTATTTTCAAACACTAAGGAGCTTTTTGACATCTGTTAAACGTCTTTTTGTTTTTTG  
TTAAGTCTTTTACATTTTAAATAGTTTTTGAAGACAATCTAGGTTAAGCAAGAGCAAAGTGCCA  
TTGTTTGCCTTTAATTGGGGGGTGGGAAGGGAAAGAGGGTACTTGCCACATAGTTTCCTTTTT  
AACTGCACTTTCTTTATATAATCGTTTGCATTTTGTACTTGCTACCCTGAGTACTTTCAGGA  
AGACTGACTTAAATATTCGGGGTGAGTAAGTAGTTGGGTATAAGATCTGAACTTTTCATCTGC  
AGAGGCAAGAAAAATATTTGACATTGTGACTTGACTGTGGAAGATGATGGTTGCATGTTTCTA  
GTTTGTATATGTTTCCATCTTTGTGATAAGATGATTTAATAAATCTCTTTAAATACTAAAAAA  
AAAAAAA

## FIGURE 4

MVSMSFKRNRSDRFYSTRCCGCCHVRTGTIILGTWYMVVNLLMAILLTVEVTHPN SMPAVNIQ  
YEVI GNYYSSE RMADNACVLF AVSVLMFI ISSMLVYGAISYQVGWLI PFFCYRLFDFVLSCLV  
AISSLTYLPRIKEYLDQLPDFPYKDDLLALDSSCLLFIVLVFFALFII FKAYLINC VWN CYKY  
INNRNVPEIAVYP AFESTSSVRFANL

### Important features of the protein:

**Transmembrane domain (Possible type II transmembrane protein):**

amino acids 30-49, 81-100, 111-131, 158-175

N-glycosylation site.

amino acids 9-13

### Tyrosine kinase phosphorylation sites.

amino acids 8-16, 193-202

N-myristoylation site.

amino acids 68-74







**FIGURE 7**

CGCGAATGAAGTTTGCATTTTCCTCTGTTCTTGAGCCCAGCTTCTTCTCGTCTCCCACCCCAG  
 CTTCCCGGCATTGGAAGAAGGGACCGTCCTCTTCCTTGTCTTGGCCACCCAAATCCTGGTATC  
 GAAAGGGTTGAACGGACCGGAAGTGTGCAGCAGCGACGGGTCCCCAGCTAATCGACGCCGGAA  
 GTAGCAATTACTAGACAAGCATTCCGCCGCCGGCTTCGCT**ATG**GCGGCAATTCCCCAGATTC  
 CTGGCAGCCACCCAACGTTTACTTGGAGACCAGCATGGGAATCATTGTGCTGGAGCTGTACTG  
 GAAGCATGCTCCAAAGACCTGTAAGAACTTTGCTGAGTTGGCTCGTCGAGGTTACTACAATGG  
 CACAAAATTCCACAGAATTATCAAAGACTTCATGATCCAAGGAGGTGACCCAACAGGGACAGG  
 TCGAGGTGGTGCATCTATCTATGGCAAACAATTTGAAGATGAACTTCATCCAGACTTGAAATT  
 CACGGGGGGCTGGAATTCTCGCAATGGCCAATGCGGGGCCAGATACCAATGGCAGCCAGTTCTT  
 TGTGACCCTCGCCCCCAGTGGCTTGACGGCAAACACACCATTTTTGGCCGAGTGTGTCA  
 GGGCATAGGAATGGTGAATCGCGTGGGAATGGTAGAAACAACTCCCAGGACCGCCCTGTGGA  
 CGACGTGAAGATCATTAAGGCATACCCTTCTGGG**TAG**ACTTGCTACCCTCTTGAGCAGCTCTT  
 CTGAGATGGCCCCAGTGAACCAGCTTCTAGATGACATAGAATGACATGTAATGCTAAATTTCA  
 TTTTGGCTTTGCAAGTCATGAAGCTTAGGAGGCCTGGCATCTTGGGTGAGTTAGAGATGGAAG  
 TACATTTTAATAGGATGCTTCTTTTCTTCCCCAGTGCCTAGGTTGCCAGAGCATTTGCAC  
 AAATGCCCCGTGTTTATCAATAGGTGACTACTTACTACACATGAACCATAATGCTGCTTCTTGT  
 GCATGTCTGCTCTGATATACGTCGAACAATGTAGCAGCCACTGTCATTTCTCAGTGGTTTTGC  
 CTAACCAAACCTTCTTCCTAAGGAGATTTATATTCTGGCCTACACAGCAGTCCTTGATGGCTGA  
 CAGCCACAGAATTCCAAACCAAGTAGTGTCTGTCAGCCCTCTTAACCTCTGTGCACGCCCTATT  
 TCAGTCTTTTACATTTGTTCTTCTAGGGAATGTATGCATCTCTATATATATTTTCCCTCTCAA  
 AACCAGAACATCAACAGTGCTGTTTCTGACACTTCAGACATCCCACGCAAAGCCACATTGAAT  
 TTTTGCCAAATGAAAAACACATCCAACAATCAAGTTTCTAAGAAGGTGTCAAGTGGGGAATAA  
 TAATAATGTATAATAATCAAGAAATTAGTTTATTAAAAGGAAGCAGAAGCATTGACCATTTTT  
 TCCCAGAGAAGAGGAGAAATCTGTAGTGAGCAAAGGACAGACCATGAATCCTCCTTGAGAAGT  
 AGTACTCTCAGAAAGGAGAAGCGCCACTCAAGTTCTTTTAACCCAAGACTTTAGAGAAATTAG  
 GTCCAAGATTTTTATATGTTTCAAGTTGTTTATGTATAAAAATAACTTTCTGGATTTTGTGGGGA  
 GGAGCAGGAGAGGAAGGAAGTTAATACCTATGTAATACATAGAACTTCCACAATAAAATGCC  
 ATTGATGGTTAAAAAAAAAAAAAAAAAAAA





**FIGURE 10**

MWHEARKHERKLRGMMVDYKKRAERRREYYEKIKKDPAQFLQVHGRACKVHLDSAVALAAESP  
 VNMPWQGDNTNMIDRFDVRAHLHDHIPDYTPPLLTTISPEQESDERKCNERYRGLVQNDFAG  
 ISEEQCLYQIYIDELYGGLQRPSEDEKKKLAEEKASIGYTYEDSTVAEVEKAAEKPEEEESAA  
 EEESNSDEDEVIPDIDVEVDVDELNQEQVADLNKQATTYGMADGDFVRMLRKDKKEAEAIKHA  
 KALEEEKAMYSGRRSRRQRREFREKRLRGRKISPPSYARRDSPTYDPYKRSPSESSSESRSRS  
 RSPTPGREEKITFITTSFGGSDEEAAAAAAAAAASGVTTGKPPAPPQPGGPAPGRNASARRRSS  
 SSSSSSSASRTSSSRSSSRSSSRSSSRGGGYRSGRHARSRSRSWSRSRSRSRRYSRSRSRGRR  
 HSGGGSRDGHRYSRSPARRGGYGPRRRSRSRSHSGDRYRRGGRGLRHHSSSRSSRSSLSPSR  
 SRSLTRSRSHSPSPSQSRRSRSRSRSQSPSPSPAREKLTRPAASPAVGEKLKKTEPAAGKETGA  
 AKVTQADASGEAETEDAEGAEQAVQGG

**Important features:****N-glycosylation site:**

amino acids 370-373

**Glycosaminoglycan attachment site:**

amino acids 443-446

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 159-162, 282-285, 291-294, 374-377, 375-378, 430-433,  
 440-443, 466-469

**Casein kinase II phosphorylation site:**

amino acids 149-152, 166-169, 171-174, 187-190, 193-196, 195-198,  
 303-306, 307-310, 335-338, 571-574

**N-myristoylation sites:**

amino acids 118-123, 229-234, 350-355, 446-451, 586-591

**Amidation sites:**

amino acids 263-266, 280-283, 438-441

**FIGURE 11**

GGTAGGCGCGCCACAGACCTGAGACGGGTTGGGACTGGGCTGCGTCACGCGCGGGCTCTAAGCG  
CCCGGGGCCCCGCCAGTGGCCGGCACAGCCAATCGCAGCGCGGGAAGGCGGTGGGGGCGGGG  
AAGGCCGCCTGGAACTTAAATCCCGAGGCGGGCGAACCTGCACCAGACCGCGGACGTCTGTA  
ATCTCAGAGGCTTGTTTTGCTGAGGGTGCCTGCGCAGCTGCGACGGCTGCTGGTTTTGAAACAT  
GAATCTTTCGCTCGTCCTGGCTGCCTTTTGCTTGGGAATAGCCTCCGCTGTTCCAAAATTTGA  
CCAAAATTTGGATACAAAGTGGTACCAGTGGAAGGCAACACACAGAAGATTATATGGCGCGAA  
TGAAGAAGGATGGAGGAGAGCAGTGTGGGAAAAGAATATGAAAATGATTGAACTGCACAATGG  
GGAATACAGCCAAGGGAAACATGGCTTCACAATGGCCATGAATGCTTTTGGTGACATGACCAA  
TGAAGAATTCAGGCAGATGATGGGTGCTTTCGAAACCAGAAATTCAGGAAGGGGAAAGTGTT  
CCGTGAGCCTCTGTTTCTTGATCTTCCCAAATCTGTGGATTGGAGAAAGAAAGGCTACGTGAC  
GCCAGTGAAGAATCAGAAACAGTGTGGTTCTTGTTGGGCTTTTAGTGCGACTGGTGCTCTTGA  
AGGACAGATGTTCCGGAAAACCTGGGAAACTTGTCTCACTGAGCGAGCAGAATCTGGTGGACTG  
TTCGCGTCCTCAAGGCAATCAGGGCTGCAATGGTGGCTTCATGGCTAGGGCCTTCCAGTATGT  
CAAGGAGAACGGAGGCCTGGACTCTGAGGAATCCTATCCATATGTAGCAGTGGATGAAATCTG  
TAAGTACAGACCTGAGAATTCTGTTGCTAATGACACTGGCTTCACAGTGGTCGCACCTGGAAA  
GGAGAAGGCCCTGATGAAAGCAGTCGCAACTGTGGGGCCCATCTCCGTTGCTATGGATGCAGG  
CCATTTCGTCTTCCAGTTCTACAAATCAGGCATTTATTTTGAACCAGACTGCAGCAGCAAAAA  
CCTGGATCATGGTGTTCTGGTGGTTGGCTACGGCTTTGAAGGAGCAAATTCGAATAACAGCAA  
GTATTGGCTCGTCAAAAACAGCTGGGGTCCAGAATGGGGCTCGAATGGCTATGTAAAAATAGC  
CAAAGACAAGAACAACCACTGTGGAATCGCCACAGCAGCCAGCTACCCCAATGTGTGAAGCTGA  
TGGATGGTGAGGAGGAAGGACTTAAGGACAGCATGTCTGGGGAAATTTTATCTTGAAACTGAC  
CAAACGCTTATTGTGTAAGATAAACCAGTTGAATCATGGAGGATCCAAGTTGAGATTTTAATT  
CTGTGACATTTTTACAAGGGTAAATGTTACCACTACTTTAATTATTGTTATACACAGCTTTA  
TGATATCAAAGACTCATTGCTTAATTCTAAGACTTTTGAATTTTCATTTTTTAAAAAGATGTA  
CAAAACAGTTTGAAATAAATTTTAATTCGTATATA

**FIGURE 12**

MNLSLVLA AFCLGIASAVPKFDQNLDTKWYQWKATHRRLYGANEEGWRRRAVWEKNMKMIELHN  
GEYSQKGHGFTMAMNAFGDMTNEEFRQMMGCFRNQKFRKGKVFREPLFLDLPKSVDWRKKGYV  
TPVKNQKQCGSCWAFSATGALEGQMFRKTGKLVSLSEQNLVDCSRPQGNQGCNGGFMARAFQY  
VKENGGLDSEESYPYVAVDEICKYRPENSVANDTGFTVVAPGKEKALMKAVATVGPISVAMDA  
GHSSFQFYKSGIYFEPDCSSKNLDHGVLVVGYGFEGANSNSKYWLVKNSWGP EWGSNGYVKI  
AKDKNNHCGIATAASYPNV

**Important features:****Signal sequence**

amino acids 1-17

**N-glycosylation sites.**

amino acids 2-6, 221-225, 292-296

**N-myristoylation sites.**

amino acids 13-19, 93-99, 136-142, 145-151, 174-180, 177-183,  
180-186, 194-200, 288-294, 324-330

**Eukaryotic thiol (cysteine) proteases cysteine active site.**

amino acids 132-144

**Eukaryotic thiol (cysteine) proteases histidine active site.**

amino acids 275-286



**FIGURE 13**

GGCGGCGTCATGTGATCCGCTTCCCTGCTCCTTTAAGCGTCCACAGGCGGCGGAGCGGCCACA  
ATCACAGCTCCGGGCATTGGGGGAACCCGAGCCGGCTGCGCCGGGGGAATCCGTGCGGGCGCC  
TTCCGTCCCGGTCCCATCCTCGCCGCGCTCCAGCACCTCTGAAGTTTTGCAGCGCCAGAAAG  
GAGGCGAGGAAGGAGGGAGTGTGTGAGAGGAGGGAGCAAAAAGCTCACCCCTAAAACATTTATT  
TCAAGGAGAAAAGAAAAAGGGGGGCGCAAAA**ATG**GCTGGGGCAATTATAGAAAACATGAGCA  
CCAAGAAGCTGTGCATTGTTGGTGGGATTCTGCTCGTGTTCCAAATCATCGCCTTTCTGGTGG  
GAGGCTTGATTGCTCCAGGGCCCCACAACGGCAGTGTCTACATGTCGGTGAAATGTGTGGATG  
CCCGTAAGAACCATCACAAGACAAAATGGTTCGTGCCTTGGGGACCCAATCATTGTGACAAGA  
TCCGAGACATTGAAGAGGCAATTCCAAGGGAAATTGAAGCCAATGACATCGTGTTTTCTGTTC  
ACATTCCCCTCCCCACATGGAGATGAGTCCTTGGTTCCAATTCATGCTGTTTATCCTGCAGC  
TGGACATTGCCTTCAAGCTAAACAACCAATCAGAGAAAATGCAGAAGTCTCCATGGACGTTT  
CCCTGGCTTACCGTGATGACGCATTTGCTGAGTGGACTGAAATGGCCCATGAAAGAGTACCAC  
GGAAACTCAAATGCACCTTCACATCTCCCAAGACTCCAGAGCATGAGGGCCGTTACTATGAAT  
GTGATGTCCTTCCCTTTCATGGAAATTGGGTCTGTGGCCATAAGTTTTACCTTTTAAACATCC  
GGCTGCCTGTGAATGAGAAGAAGAAAATCAATGTGGGAATTGGGGAGATAAAGGATATCCGGT  
TGGTGGGGATCCACCAAAATGGAGGCTTACCAAGGTGTGGTTGCCATGAAGACCTTCCTTA  
CGCCAGCATCTTCATCATTATGGTGTGGTATTGGAGGAGGATCACCATGATGTCCCGACCCC  
CAGTGCTTCTGAAAAAGTCATCTTTGCCCTTGGGATTTCCATGACCTTTATCAATATCCCAG  
TGGAAATGGTTTTCCATCGGGTTTGAAGTGGACCTGGATGCTGCTGTTTGGTGACATCCGACAGG  
GCATCTTCTATGCGATGCTTCTGTCTTCTGGATCATCTTCTGTGGCGAGCACATGATGGATC  
AGCACGAGCGGAACCACATCGCAGGGTATTGGAAGCAAGTCGGACCCATTGCCGTTGGCTCCT  
TCTGCCTCTTCATATTTGACATGTGTGAGAGAGGGGTACAACCTCACGAATCCCTTCTACAGTA  
TCTGGACTACAGACATTGGAACAGAGCTGGCCATGGCCTTCATCATCGTGGCTGGAATCTGCC  
TCTGCCTCTACTTCTGTCTTCTATGCTTCATGGTATTTTCAAGGTGTTTCGGAACATCAGTGGGA  
AGCAGTCCAGCCTGCCAGCTATGAGCAAAGTCCGGCGGCTACACTATGAGGGGGCTAATTTTTTA  
GGTTCAAGTTCCCTCATGCTTATCACCTTGGCCTGCGCTGCCATGACTGTCATCTTCTTCATCG  
TTAGTCAGGTAACGGAAGGCCATTGGAAATGGGGCGGCGTCACAGTCCAAGTGAACAGTGCCT  
TTTTTACAGGCATCTATGGGATGTGGAATCTGTATGTCTTTGCTCTGATGTTCTTGTATGCAC  
CATCCCATAAAAACTATGGAGAAGACCAGTCCAATGGCGATCTGGGTGTCCATAGTGGGGAAG  
AACTCCAGCTCACCACCCTATCACCCATGTGGACGGACCCACTGAGATCTACAAGTTGACCC  
GCAAGGAGGCCCAGGAG**TAGG**AGGCTGCAGCGCCCGGCTGGGACGGTCTCTCCATACCCCAGC  
CCCTCTAACTAGAGTGGGGAGCATGCCAGAGAGAGCTCAATGTACAAATGAATGCCTCATGGC  
TCTTAGCTGTGGTTTTCTTGGACCAGCGGCATGGACATTTGTGAGTTTGCCTTCTGACGGTAGC  
TTTTGGAGGAAGATTCCCTGCAGCCACTAATGCATTGTGTATGATAACAAAACTCTGGTATGA  
CACATTTTCTGTGATCATTGTTAATTAGTGACATAGTAACATCTGTAGCAGCTGGTTAGTAAA  
CCTCATGTGGGGGTGGGGTGGGGTGTATTCCTTGGGGGATGGTTTGGGCCGAATGGGGAGTG  
GAATATTTGACATTTTTCCTGTTTTAAATTCTAGGATAGATTTTAAACATCCTTTGCGGTCCCA  
GTCCAAGGTAGGCTGGTGTCATAGTCTTCTCACTCCTAATCCATGACCACTGTTTTTTTCTTA  
TTTATATCACCAGGTAGCCTACTGAGTTAATATTTAAGTTGTCAATAGATAAGTGTCCCTGTT  
TTGTGGCATAATATAACTGAATTTTCATGAGAAGATTTTATCCACCAGGGGTATTTACGCTTTG  
AAACCAATCTGTGTATCTAATACTAACCAATCTGTTGGATGTGGATTTTAAAAAATGTTTGC  
TAACTACCCAAGTAAGATTTTACTGTATTAAATGGCCTTCGGGTCTGAAAAGCTTTTTTTAAACC  
TCTTGCTTAAATGCGTTTTATTTTGATAAGATACTTCAAATAGCCTCCAAAAGTGTAGATCC  
AATCACTTAAATAAACCTGTATGTATATGCAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 14**

MAGAI IENMSTKKLCIVGGILLVFQIIAFLVGGGLIAPGPTTAVSYMSVKCVDARKNHHKTKWF  
VPWGPNHCDKIRDIEEAIPREIEANDIVFSVHIPLPHMEMSPWFQFMLFILQLDIAFKLNNQI  
RENAEVSMDVSLAYRDDAFAEWTEMAHERVPRKLKCTFTSPKTPEHEGRYYECDVLPFMEIGS  
VAHKFYLLNIRLPVNEKKKINVGIGEIKDIRLVGIHQNGGFTKVWFAMKTFLTPSIFII MVWY  
WRRITMMSRPPVLLLEKVI FALGISMTFINIPVEWFSIGFDWTWMLLFGDIRQGI FYAMLLSFW  
IIFCGEHMMDQHERNHIAGYWKQVGPIAVGSFCLFI FDMCERGVQLTNPFYSIWTTDIGTELA  
MAFIIVAGICLCLYFLFLCFMVFQVFRNISGKQSSLPAMSKVRRLLHYEGLIFRKFFLMLITLA  
CAAMTVIFFIVSQVTEGHWKWGGVTVQVNSAFFTGIYGMWNLYVFALMFLYAPSHKNYGEDQS  
NGDLGVHSGEELQLTTTITHVDGPTEIYKLTRKEAQE

**Important features of the protein:****Signal peptide:**

amino acids 1-42

**Transmembrane domains:**

amino acids 239-253, 269-284, 302-318, 338-352, 377-399, 434-452,  
471-488

**N-glycosylation sites.**

amino acids 8-12, 406-410

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 254-258

**N-myristoylation sites.**

amino acids 223-229, 274-280, 305-311, 358-364, 374-380, 386-392,  
509-515



**FIGURE 16**

MEESPLSRAPSRGGVNFLNVARTYIPNTKVECHYTLPPGTMPASDWIGIFKVEAACVRDYHT  
 FVWSSVPESTTDGSPIHSTSVQFQASYLPKPGAQLYQFRYVNRQGQVCGQSPPFQFREPRPMDE  
 LVTLEEADGGSDILLVVPKATVLQNQLDESQQERNDLMQLKLQLEGQVTELRSRVQELERALA  
 TARQEHTELMEQYKGISRSHGEITEERDILSRQQGDHVARILELEDDIQTISEKVLTKVELD  
 RLRDTVKALTREQEKLKGQLKEVQADKEQSEAEQLQVAQQENHHLNLDLKEAKSWQEEQSAQAQ  
 RLKDKVAQMKDTLGQAQQRVAELEPLKEQLRGAQELAASSQQKATLLGEELASAAAARDRTIA  
 ELHRSRLEVAEVNGRLELGLHLKEEKCQWSKERAGLLQSVAEKDKILKLSAEILRLEKAVQ  
 EERTQNQVFKTELAREKDSSLVQLSESKRELTELRSAALRVLQKEKEQLQEEKQELLEMYMRKLE  
 ARLEKVADEKWNEDATTEDEEEAAVGLSCPAALTDSEDESPEDMRLPPYGLCERGDPGSSPAGP  
 REASPLVVISQPAPISPHLSGPAEDSSSDSEAEDEKSVLMAAVQSGGEEANLLLPELGSAFYD  
 MASGFTVGTLSSETSTGGPATPTWKECPICKERFPAESDKDALEDHMDGHHFFSTQDPFTFE

**Important features:****Casein kinase II phosphorylation sites:**

amino acids 28-31, 43-46, 68-71, 72-75, 129-132, 156-159, 208-  
 211, 239-242, 282-285, 305-308, 376-379, 383-383, 468-471, 520-  
 523, 521-524, 537-540, 539-542, 543-546, 593-596, 595-598, 597-  
 600, 612-615, 639-642, 652-655, 667-670, 683-686

**N-myristoylation sites:**

amino acids 39-44, 107-112, 204-209, 414-419, 561-566, 613-618

**Cell attachment sequence:**

amino acids 557-559

**Leucine zipper pattern sequence:**

amino acids 163-184, 475-496, 482-503

**FIGURE 17**

GCAAGTTGGGAATTTTAGACTGTCACTGCACATGGACCTCTGGGAAGACGTCTGGCGAGAGCT  
AGGCCCCTGGCCCTACAGACGGATCTTGCTGGCTCACCTGTCCCTGTGGAGGTTCCCCTGGG  
AAGGCAAG**ATG**CCCCAACACAGCACTGCTCTGTCAATTGGCCAATGTTACCTACATCACCATGG  
AAATTTTCATTGGACTCTGCGCCATAGTGGGCAACGTGCTGGTCATCTGCGTGGTCAAGCTGA  
ACCCAGCCTGCAGACCACCACCTTCTATTTTCATTGTCTCTCTAGCCCTGGCTGACATTGCTG  
TTGGGGTGCTGGTCATGCCTTTGGCCATTGTTGTCAGCCTGGGCATCACAATCCACTTCTACA  
GCTGCCTTTTTATGACTTGCCCTACTGCTTATCTTTACCCACGCCTCCATCATGTCTTGCTGG  
CCATCGCTGTGGACCGATACTTGCGGGTCAAGCTTACCGTCAGATTGAGAATTCCTGGGCTCC  
CTGGGTGCATTCTATCATTCCAGTTGAAAGTTTGCTTCCTTCCAGTCATGTGGCTCTTCATTC  
TACTCTCCTTGGCTCTCATTTGAGATGCCATGGTCATGGATGAAAAGGTCAAGAGAAGCTTTG  
TGCTGGACACGGCTTCTGCCATCTGCAACTACAATGCCACTACAAGAATCACCCCAAATACT  
GGTGCCGAGGCTATTTCCGTGACTACTGCAACATCATCGCCTTCTCCCCTAACAGCACCAATC  
ATGTGGCCCTGAGGGACACAGGGAACCAGCTCATTGTCACTATGTCCTGCCTGACCAAAGAGG  
ACACGGGCTGGTACTGGTGTGGCATCCAGCGGGACTTTGCCAGGGATGACATGGATTTTACAG  
AGCTGATTGTAAGTACGACAAAGGAACCCTGGCCAATGACTTTTGGTCTGGGAAAGACCTAT  
CAGGCAACAAAACCAGAAGCTGCAAGGCTCCCAAAGTTGTCCGCAAGGCTGACCGCTCCAGGA  
CGTCCATTCTCATCATTTGCATACTGATCACGGGTTTGGGAATCATCTCTGTAATCAGTCATT  
TGACCAAAAGGAGGAGAAGTCAAAGGAATAGAAGGGTAGGCAACACTTTGAAGCCCTTCTCGC  
GTGTCCTGACTCCAAAGGAAATGGCTCCTACTGAACAGATG**TGA**CTGAAGATTTTTTTTAATTT  
AGTTCATAAAGTGATGCTACAACAGAATAATCACCATGACAACTGGCCACACCTCAGAGACT  
GATTCTGATCTCCAGGAATTCTGAAGGACCCTCTATCCTTGACAACAATCATTTGCAGCCAG  
GTAGCAACGGCGGTAGTCAGAGGAGCTATGATAGACCACACCAAGCAAGGCTGCCCTCAAAT  
AACATCTCAAGATCTTAGTTCTTATGCATTCCATCAGTCAGAAGTGAAGAAGAGGTGGAGAAT  
CTGGATTGGGGACCAGGAAATCACTTGTATTTTGTAGCCAATAAATTCCTAGCCAGTGTTGA  
ATGAAAAAAAAAAAAA

## **FIGURE 18**

MPNNSTALSLANVTYITMEIFIGLCAIVGNVLVICVVKLNPSLQTTTFYFIVSLALADIAVG  
 LVMPPLAIVVSLGITIHFYSCLFMTCLLLIFTHASIMSLLAIAVDRYLRVKLTVRFRIPGLPGC  
 ILSFQLKVCFLPVMWLFILLSLALISDAMVMDEKVKRSFVLDTASAICNYNHYKNHPKYWCR  
 GYFRDYCNIIAFSPNSTNHVALRDTGNQLIVTMSCLTKEDTGWYWCGIQRDFARDDMDFTTELI  
 VTDDKGTLANDFWSGKDLSGNKTRSCAPKVVVRKADRSRTSILIICILITGLGIISVISHLTK  
 RRRSQNRNRVGNLTKPFSRVLTPKEMAPTEQM

### **Important features of the protein:**

#### **Transmembrane domains:**

amino acids 16-35, 62-80, 89-101, 134-152, 292-311

#### **N-glycosylation sites.**

amino acids 3-7, 4-8, 12-16, 204-208, 273-277

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 316-320

#### **N-myristoylation sites.**

amino acids 122-128, 125-131, 258-264

#### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 214-225

#### **G-protein coupled receptors proteins.**

amino acids 29-59, 76-116

**FIGURE 19**

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCCTGCGATTGAGCTGCGGGTCGCGGCCGGCGCCGGCCCTCTCCAATG  
GCAAATGTGTGTGGCTGGAGGCGAGCGGAGGCTTTCGGCAAAGGCAGTCGAGTGTGTCAGACCGGGGGCGAGTC  
CTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCTATTACGAGGGGAGCGCCCGGGGGGCTGTGCGACT  
CCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGGAGAAGAAGAAAAGCGGAAAAGAGGCAGATTCACGTG  
TTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCCCCCTG  
GTTTGTGTGTTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCCTGGGCG  
AATCCCACATCTGTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTGCAATCTGCGAGTGAAGAGGG  
ACGAGGGAAAAGAAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCCAAAAGAAGCACCAGATCAGCAAAA  
AAAGAAGATGGGCCCCCGAGCCTCGTGCTGTGCTTGTGTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTC  
GGCCTTCCTGTGCGACCAACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCAACATCATCCT  
GGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGCATCATGGAGCA  
GGGCGGGGCGCACTTCATCAACGCCCTTCGTGACCACACCCATGTGCTGCCCCCTCACGCTCCTCCATCCTCACTGG  
CAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAAGTGTCTCCTCGCCCTCCTGGCAGGCACAGCACGA  
GAGCCGCACTTTGCCGTGTACCTCAATAGCACTGGCTACCCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATA  
CAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTGCGACTCCTTAAAACTCCCGCTTTTATAACTACAC  
GCTGTGTGCGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCACAGACCTCATCACAA  
TGACAGCGTGAGCTTCTTCCGCGACGTCCAAGAAGATGTACCCGCGACAGGCCAGTCCCTCATGGTCATAGCCATGC  
AGCCCCCAGCGCCCTGAGGATTGAGCCCCACAATATTACGCCTCTTCCCAAACGCATCTCAGCACATCACGCC  
GAGCTACAACTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACAT  
GGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCGAGACCCCTCATGTGCGTGGACGACTCCATGGAGACGATTTA  
CAACATGTGTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGCCGACCACGGTTACCACATCGG  
CCAGTTTGCCCTGGTGAAAGGGGAAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTTACGTGAGGGGGCC  
CAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTGGCCCCCACCATCCTGGACATTGC  
AGGCTTGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCTGGACACGGAGCGGCCGGTGAATCG  
GTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGGAGAGAGGCAAGCTGCTACACAAGAG  
AGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTACCAGCGTGTGAAGGACCTGTGTGACGG  
TGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTGTGTGGAGGACGCCACGGGGAAGCTGAA  
GCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCCTCTCCAACCTCGTGCCCAAGTACTACGG  
GCAGGGCAGCGAGGCTGCACCTGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGGACGCCGGAAAAAACTCTT  
CAAGAAGAAGTACAAGGCCAGCTATGTCCGAGTCTGCTCCATCCGCTCAGTGGCCATCGAGGTGGACGGCAGGGT  
GTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCACCAAGCGGCACTGGCCAGGGGGCCCTGAGGA  
CCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGCCTTCCCGACTACTCAGCCGCCAACCCCATTA  
AGTGCACATCGGTGCTACATCCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGC  
CTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAACCTGCAGAACAAAATTAAGAACCTGAGGGGAAGT  
CCGAGGTACCTGAAGAAAAAGCGGCCAGAAGAATGTGACTGTACAAAAATCAGCTACCACACCCAGCACAAAAGG  
CCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAGGAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTT  
GCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGCTCAAGCGCTGCAGAACAACGACACGTGCAGCATGCC  
AGGCTCACGTGCTTCACCACGACAACCAGCACTGGCAGACGGCGCCTTCTGGACACTGGGGCCTTCTGTGTC  
CTGCACCAGCGCCAACAATAACACGTACTGGTGCTGAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGA  
ATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACT  
GGACAGGGATGTCTCAACCAGCTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAA  
CCCCCGGACTCGAAACATGGACCTGGATGGAGGAAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCC  
AGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACCTGTGGGAAGGCTGGGAAGGT**TAA**GAAACAAACAGAGGT  
GGACCTCCAAAAACATAGAGGCATCACCTGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACC  
TGTGCTATTGGCCAGGAGGCCTGAGAAAGCAAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCA  
GCAGGAGCAGAGATAACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCAC  
AAAATGCATTTTTTTCGTATCAAAAAGTCACCACTAACCCCTCCCCAGAAGCTCACAAAGGAAAAACGGAGAGAGCG  
AGCGAGAGAGATTTCTTGGAAATTTCTCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAAGCA  
GTCCTGTTCTAAATCCTCTTATTCTTTTGGTTTGTACAAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGG  
AGAGGCTGAAAACAGTGCAGAGACGTTTGACAATTGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTAT  
AAACCTGGTTGCCTCTGAAGAACTGCCTTCAATTGATATATGTGACTATTTACATGTAATCAACATGGGAACCT  
TTTAGGGGAACCTAATAAGAAATCCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAA  
GAAAAA

**FIGURE 20**

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRRRNIRPNIILVLTDDQDVELGSMQV  
 MNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHESTR  
 FAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSDYSK  
 DYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITPSYNY  
 APNPDKHWIMRYTGPMKPIHMEFTNMLQQRKLQTLMSVDDSMETIYNMLVETGELDNTYIVYT  
 ADHGYHIGQFGLVKGKSMPIYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDIAGLDIP  
 ADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEENFLPKYQR  
 VKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKYYGQGSEAC  
 TCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVAI EVDGRVYHVGLGDAAQPRNLTKRHW  
 PGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLH  
 IDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSLHPFRKGLQEKD  
 KVWLLREQRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNT  
 YWCMRTINETHNFLCFEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGY  
 KQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

**Important features:****Signal peptide:**

amino acids 1-17

**Sulfatases signature 1.**

amino acids 86-99

**Homologous region to sulfatase:**

amino acids 87-106, 133-146, 216-229, 291-320, 365-375

**N-glycosylation sites.**

amino acids 65-69, 112-116, 132-136, 149-153, 171-175, 198-202,  
 241-245, 561-565, 608-612, 717-721, 754-758, 764-768



**FIGURE 21**

GGGCGCGCGAGAGCTGCTAGGGCGGTTTCTCTGCCTCGGGCCTGTTGGGCAGGGCCGGCT  
AAGGTGCGCGTGCTCGCTGGTTCTAACCCTTCTGTTGGGCGTTTCTGCTGAGAGGCGGGA  
GGCGCTGAGAGTCTGTGCGGAGGTCCGTGGACAGACTGCTTTGCTCGTTGTTGCTCTTCG  
GAGGCGGCGATCCCCGAAGGCGAGCTGAAATACGGCTGCAGGCTACAATTTGCAGCCGAC  
GATTATGGAAGACGGAAGCGGGAGAGGTGGCCCACCCTC**ATG**GAGCGCTTGTGCTCGGAT  
GGCTTCGCATTTCCCCAATACCCCATTAACCGTATCATCTGAAGAGGATCCACAGAGCT  
GTCTTACATGGTAATCTAGAGAACTGAAGTACCTTCTGCTCACGTATTATGACGCCAAT  
AAGAGAGACAGGAAGGAAAGGACCGCCCTACATTTGGCCTGTGCCACTGGCCAACCGGAA  
ATGGTACATCTCCTGGTGTCCAGAAGATGTGAGCTTAACCTCTGCGACCGTGAAGACAGG  
ACACCTCTGATCAAGGCTGTACAACTGAGGCAGGAGGCTTGTGCAACTCTTCTGCTGCAA  
AATGGCGCCAATCCAAATATTACGGATTTCTTTGGAAGGACTGCTCTGCACTACGCTGTG  
TATAATGAAGATACATCCATGATAGAAAACTTCTTTCACATGGTACAAATATTGAAGAA  
TGCAGCAAGGTA**TAG**GTCAACCAATGTTATTTTCAAACCTATCTGAAATGAATTTATTTTA  
ACATTGACACATGTAAGGGTCAATTTTTCATATTTGGAAGCTCAAACATTCCTTGAATGA  
AAATATTTTGAAATGCCTTAACTGTCTAAGATTTTACTTTAAATATTGGAACCTTTTAAAG  
AAGCATTATAGGGAACAGCCTTTTTTTCATGCACTTATGGTAAATAACTATAAAAACAAAT  
GAATTACAATAAATTTATAATTCATGACAACTGAATTTGGGAAAGGTAATAGTTAAGTGT  
TTTTCCACTAAATTACTTTTT

## **FIGURE 22**

MERLCSDGFAFPQYPIKPYHLKRIHRAVLHGNLEKLYLLLTYYDANKRDRKERTALHLACAT  
GQPEMVHLLVSRRCENLDCDREDRTPLIKAVQLRQEACATLLLQNGANPNITDFFGRTALHYA  
VYNEDTSMIEKLLSHGTNIEECSKV

### **Important features of the protein:**

#### **N-glycosylation site.**

amino acids 113-117

#### **N-myristoylation site.**

amino acids 109-115

#### **Microbodies C-terminal targeting signal.**

amino acids 149-153

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 23**

GAGGCAGAAAGGCAGAAAGGAGAAAATTCAGGATAACTCTCCTGAGGGGTGAGCCAAGCCCTG  
CCATGTAGTGCACGCAGGACATCAACAAACACAGATAACAGGAAATGATCCATTCCTGTGGT  
CACTTATTCTAAAGGCCCCAACCTTCAAAGTTCAAGTAGTGAT**ATG**GATGACTCCACAGAAAG  
GGAGCAGTCACGCCTTACTTCTTGCCTTAAGAAAAGAGAAGAAATGAAACTGAAGGAGTGTGT  
TTCCATCCTCCCACGGAAGGAAAGCCCCTCTGTCCGATCCTCCAAAGACGGAAAGCTGCTGGC  
TGCAACCTTGCTGCTGGCACTGCTGTCTTGCTGCCTCACGGTGGTGTCTTTCTACCAGGTGGC  
CGCCCTGCAAGGGGACCTGGCCAGCCTCCGGGCAGAGCTGCAGGGCCACCACGCGGAGAAGCT  
GCCAGCAGGAGCAGGAGCCCCCAAGGCCGGCCTGGAGGAAGCTCCAGCTGTCACCGCGGGACT  
GAAAATCTTTGAACCACCAGCTCCAGGAGAAGGCAACTCCAGTCAGAACAGCAGAAATAAGCG  
TGCCGTTCAAGGTCCAGAAGAAACAGTCACTCAAGACTGCTTGCAACTGATTGCAGACAGTGA  
AACACCAACTATACAAAAAGGATCTTACACATTTGTTCCATGGCTTCTCAGCTTTAAAAGGGG  
AAGTGCCCTAGAAGAAAAAGAGAATAAAATATTGGTCAAAGAAACTGGTTACTTTTTTATATA  
TGGTCAGGTTTTATATACTGATAAGACCTACGCCATGGGACATCTAATTCAGAGGAAGAAGGT  
CCATGTCTTTGGGGATGAATTGAGTCTGGTGACTTTGTTTCGATGTATTCAAATATGCCTGA  
AACACTACCCAATAATTCCTGCTATTTCAGCTGGCATTGCAAACTGGAAGAAGGAGATGAACT  
CCAACCTGCAATACCAAGAGAAAATGCACAAATATCACTGGATGGAGATGTCACATTTTTTGG  
TGCATTGAACTGCTG**TGA**CCCTACTTACACCATGTCTGTAGCTATTTTCCTCCCTTTCTCTGT  
ACCTCTAAGAAGAAAGAATCTAACTGAAAATACCAAAAAAAAAAAAAAAAAA

## FIGURE 24

MDDSTEREQSRLTSCCLKREEMKLKECVSILPRKESPSVRSSKDGKLLAATLLLALLSCCLTV  
VSFYQVAALQGD LASLRAELQGHHA EKLPAGAGAPKAGLEEAPAVTAGLKI FEPPAPGEGNSS  
QNSRNKRAVQGPEETVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKRGSALEEKENKILVKE  
TGYFFIYGQVLYTDKTYAMGHLIQRKKVHVFGDELSLVT LFRCIQNMPETLPNNSCYSAGIAK  
LEEGDELQLAIPRENAQISLDGDVTTFFGALKLL

Transmembrane domain:

amino acids 47-72

N-glycosylation site.

amino acids 124-127, 242-245

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36, 173-176

N-myristoylation site.

amino acids 96-101

TNF family proteins.

amino acids 172-206



## **FIGURE 26**

MEARLRSDVIHAPLPSPVDKVAANTPSMYSQELFQLSQYLQEALHREQMLEQKLATLQRL LAI  
TQEASDTSWQALIDEDRLLSRLEVMGNQLQACSKNQTEDSLRKELIALQEDKHNYETTAKESL  
RRVLQEKIEVVRKLSEVERSLSNTEDECTHLKEMNERTQEELRELANKYNGAVNEIKDLSDKL  
KVAEGKQEEIQQKGQAEKKELQHKIDEMEEKEQELQAKIEALQADNDFTNERLTALQVRLEHL  
QEKTLKECSSLADRRRRASNQSGRRNKAFKR FVFCFSMFFDSSFG

### **Important features of the protein:**

#### **N-glycosylation sites.**

amino acids 98-102, 271-275

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 138-142, 267-271

#### **Amidation site.**

amino acids 273-277

#### **Tropomyosins proteins.**

amino acids 169-217



**FIGURE 28**

MEEKYGGDVLAGPGGGGGLGPVDVPSARLTKYIVLLCFTKFLKAVGLFESYDLLKAVHIVQFI  
 FILKLGTAFPMVLFQKPFSSGKTITKHQWIKIFKHAVAGCIISLLWFFGLTLCGPLRTLLLFE  
 HSDIVVISLLSVLFTSSGGGPAKTRGAFFIIAIVICLLLFNDNDLMAKMAEHPEGHHDSALTH  
 MLYTAIAFLGVADHKGGVLLLVLALCCKVGFHTASRKLSVDVGGAKRLQALSHLVSVLLCPW  
 VIVLSVTTESKVESWFSLIMPFATVIFVVMILDFYVDSICSVKMEVSKCARYGSFPIFISALL  
 FGNFWTHPITDQLRAMNKAHQESTEHVLSGGVVVSAIFFILSANILSSPSKRGQKGTIGYS  
 PEGTPLYNFMGDAFQHSQSIPRFIKESLKQILEESDSRQIFYFLCLNLLFTFVELFYGVLTN  
 SLGLISDGFHMLFDCSALVMGLFAALMSRWKATRIFSYGYGRIEILSGFINGLFLIVIAFFVF  
 MESVARLIDPPELDTHMLTPVSVGGLIVNLIGICAFSHAHSHAHGASQGSCHSSDHSHSHMH  
 GHSDHGHGSHGSAGGGMNANMRGVFLHVLADTLGSIGVIVSTVLIEQFGWFIADPLCSLSTA  
 ILIFLSVPLIKDACQVLLRLPPEYEKELHIALEKIQKIEGLISYRDPHFWRHSASIVAGTI  
 HIQVTSVDLEQRIVQQVTGILKDAGVNNLTIQVEKEAYFQHMSGLSTGFHDVLAMTKQMESMK  
 YCKDGTIIM

**Important features of the protein:****Signal peptide:**

amino acids 1-46

**Transmembrane domains:**

amino acids 59-77, 101-119, 150-167, 205-223, 239-258, 267-284,  
 305-324, 343-360, 421-440, 452-469, 486-505, 522-539, 592-612,  
 621-641

**N-glycosylation site.**

amino acids 721-725

**Glycosaminoglycan attachment site.**

amino acids 143-147

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 225-229

**Tyrosine kinase phosphorylation sites.**

amino acids 750-758, 756-764

**N-myristoylation sites.**

amino acids 14-20, 46-52, 102-108, 112-118, 144-150, 317-323,  
 347-353, 369-375, 372-378, 437-443, 462-468, 529-535, 549-555,  
 553-559, 579-585, 582-588, 583-589, 584-590, 605-611, 737-743

**Multicopper oxidases protein:**

amino acids 561-569



### FIGURE 29

GGCACGAGGGCAGGATATTAGAAATGGCTACTCCCCAGTCAATTTTTCATCTTTGCAATCTGCA  
TTTAAATGATAACAGAATTAATTCTGGCCTCAAAAAGCTACTATGATATCTTAGGTGTGCCAA  
AATCGGCATCAGAGCGCCAAATCAAGAAGGCCTTTCACAAGTTGGCCATGAAGTACCACCCTG  
ACAAAATAAGAGCCCGGATGCTGAAGCAAAATTTCAGAGAGATTGCAGAAGCATATGAAACAC  
TCTCAGATGCTAATAGACGAAAAGAGTATGATACACTTGGACACAGTGCTTTTACTAGTGGTA  
AAGGACAAAGAGGTAGTGGAAGTTCTTTTGAGCAGTCATTTAACTTCAATTTTGATGACTTAT  
TTAAAGACTTTGGCTTTTTTTGGTCAAAACCAAAACACTGGATCCAAGAAGCGTTTTGAAAATC  
ATTTCCAGACACGCCAGGATGGTGGTTCAGTAGACAAAGGCATCATTTCCAAGAATTTTCTT  
TTGGAGGTGGATTATTTGATGACATGTTTGAAGATATGGAGAAAATGTTTTCTTTTAGTGGTT  
TTGACTCTACCAATCAGCATACAGTACAGACTGAAAATAGATTTTCATGGATCTAGCAAGCACT  
GCAGGACTGTCACTCAACGAAGAGGAAATATGGTTACTACATACACTGACTGTTTCAGGACAGTAG  
AGTTCTTATTCTATTCTCACTAAATCCAACCTGGTTGACTCTTCCTCATTATCTTTGATGCTAA  
ACAATTTTCTGTGAACTATTTTGACAAGTGCATGATTTCACTTTAAACAATTTGATATAGCTA  
TTAAATATATTTAAGGGTTTTTTTTTTTTTGACAAATTCAACATTCAACGAGTAGACAAAATGCT  
AATTATTTCCCTGATTAGGAAAGTTTCTTTAAAAAACACGTAATTTTGCCTAGTGCTTTTTTCT  
CTACCTGCCCTTGGGCTCACTAATATCACCAGTATTATTACCAAGAAAATATTGAGTTTACCT  
GATTAACTTTTAAAAGTTAATTGTAGATTTAAATTGTGTGAACCTAATGATTTTTGCAGTGAA  
ACCTTTACTAATTCAAAGTTGCATGTTCTATGACATCTGTGACTTGCGTTGCAGAGTGACAT  
GAAACTGTATAATTGAGTCATTCAGTAAAGGAGAACAGTATCTTGGTTAATTGCTACTGAAAG  
GTTGAGAAAGGAATGGTTTGATATTTACCACAGCGCTGTGCCCTTTCTACAGTAGAACTGGGGT  
AAAGGAAATGGTTTTATTGCCCATAGTCATTTAGGCTGGAAAAAAGTTGAAAACCTAACGAAA  
TATTGCCAAGAGATTGTTATGTGTTTGGTTCAGCCTAAAAATGATTTTGTAGTGTTGAAATC  
ATAGCTACTTACATAGCTTTTTTCATATTTCTTCTTAGTTGTTGGCACTCTTAGGTCTTAGTA  
TGGATTTATGTGTTTGTGTGTGTGTAGTTTATCCTCTCTCATCTTTATCTAGAGATTGACT  
GATACCTCATTCTGTTTGTAACCAGCCAGTAATTTCTGTGCAACCTTACTATGTGCAATAT  
TTTTAAATCCTGAGAAATGTGTGCTTTTGTTTTCGGATAGACTTATTTCTTTAGTTCTGCACT  
TTTCCACATTATACTCCATATGAGTATTAATCCTATGGATACATATTAACAAGTGTCTCAT

## **FIGURE 30**

MATPQSIFIFAICILMITELILASKSYDILGVPKSASERQIKKAFHKLAMKYHPDKNKSPDA  
EAKFREIAEAYETLS DANRRKEYDTLGHSAFTSGKGQRGSGSSFEQSFNFNFDDL FKDFGFFG  
QNQNTGSKKRFENHFQTRQDGGSSRQRHHFQEFSFGGGLFDDMFEDMEKMFSFSGFDSTNQHT  
VQTENRFHGSSKHCRTVTQRRGNMVTTYTD CSGQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-23

#### **Nt-dnaJ domain signature.**

amino acids 27-59, 66-90

#### **Glycosaminoglycan attachment site.**

amino acids 96-100

#### **N-myristoylation sites.**

amino acids 32-38, 99-105, 102-108, 126-132, 211-217

FIGURE 31

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTGG  
GCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCGGGCTCTAGAACAAT  
TCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCTGAG  
ATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCAA**ATG**  
CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCTACGCA  
TTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTCTCTGTA  
CTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGAAACAGTG  
TACTATTCTGTGCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCTGGATCCCC  
AGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATCACGGCCACT  
GTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTGGAGCATCCTG  
AAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGATCACCAAAGAT  
GGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTCCTTGTGGCCTAC  
TGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGGGGGTATTCCAGTG  
CACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGACATTCGTGAAGGCC  
ATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAAGGAGAGGCCATTCCC  
CTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGTGGTCGTGCCACTGTTC  
GTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGGTGGTCCCTCCAGACACC  
TTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGGGAGGAGGTGGATGCCTGT  
GCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGATCTCA**TAG**GTTTGC~~CG~~GAAGG  
GCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACCATGAGGGGGACAAGTTGTGTT  
TCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGAGCCTGTTGTCTACAAGTCTAG  
AAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACACTGACTGAGGCTTAGGGGATGTG  
ACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACCCTGGGAAAAGTGACTTCATCCCT  
TCGGTCCTAAGTTTTTCTCATCTGTAATGGGGGAATTACCTACACACCTGCTAAACACACACAC  
ACAGAGTCTCTCTCTATATATACACACGTACACATAAATACACCCAGCACTTGCAAGGCTAGA  
GGGAAACTGGTGACACTCTACAGTCTGACTGATTCAAGTGTCTGGAGAGCCACTTTC~~CC~~CAGAAT  
TATGATGAGAATGATCAAGGACTCTACACACTGGGTGGCTTGGAGAGCCACTTTC~~CC~~CAGAAT  
AATCCTTGAGAGAAAAGGAATCATGGGAGCAATGGTGTGAGTTCACTTCAAGCCCAATGCCG  
GTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGTAGGTGACCTGGAGGAAGGTCACAGCCACA  
CTGAAAATGGGATGTGCATGAACACGGAGGATCCATGAACTACTGTAAAGTGTGACAGTGTG  
TGCACACTGCAGACAGCAGGTGAAATGTATGTGTGCAATGCGACGAGAATGCAGAAGTCAGTA  
ACATGTGCATGTTTGTGTGCTCCTTTTTTCTGTTGGTAAAGTACAGAATTCAGCAAATAAAA  
AGGGCCACCCTGGCCAAAAGCGGTAAAAA

## **FIGURE 32**

MQTFTMVLEEIWTSLEFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGET  
VYYSVEYQGEYESLYTSHIWIPSSWSCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAWSI  
LKHPFNRNSTILTRPGMEITKDG FHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSGGIP  
VHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECEVQGEAIPVLALFAFVGFM LILVVVPL  
FVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

### **Important features:**

#### **Signal peptide:**

amino acids 1-29

#### **Transmembrane domain:**

amino acids 230-255

#### **N-glycosylation sites.**

amino acids 40-44, 134-138

#### **Tissue factor proteins.**

amino acids 92-120

#### **Integrins alpha chain proteins.**

amino acids 232-263

**FIGURE 33**

GAGACACGCGAGCGGGGAGACCTCCAAGGCAGCGAGGCATCGGACATGTGTCAGCACATCTGG  
 GGCGCACATCCGTCGAGCCCGAGGGGAGATTTGCCGGAACAATTCAAACCTGCGATATTGATCT  
 TGGGGGTGACTGTCCCTGGCCGGCTGTGCGGTGGGAGTGCAGAGTGTGCACTCGCTCGGAAGTG  
 TGTGCGAGTGTGTATGTGTGTGTGCCGTGTGCGGCTCCCCCCTTCCCCCGTTTCCCGTCGA  
 GTGATGCACTTGAATGAGAATCAGAGG**GATG**GAAATAGTCTGGGAGGTGCTTTTTCTTCTTCA  
 AGCCAATTTTCATCGTCTGCATATCAGCTCAACAGAATTCACCAAAAATCCATGAAGGCTGGTG  
 GGCATACAAGGAGGTGGTCCAGGGAAGCTTTGTTCCAGTTCCTTCTTTCTGGGGATTGGTGAA  
 CTCAGCTTGAATCTTTGCTCTGTGGGGAAACGGCAGTCGCCAGTCAACATAGAGACCAGTCA  
 CATGATCTTCGACCCCTTTCTGACACCTCTTCGCATCAACACGGGGGGCAGGAAGGTCAGTGG  
 GACCATGTACAACACTGGAAGACACGTATCCCTTCGCTGGACAAGGAGCACTTGGTCAACAT  
 ATCTGGAGGGCCCATGACATACAGCCACCGGCTGGAGGAGATCCGACTACACTTTGGGAGTGA  
 GGACAGCCAAGGGTCGGAGCACCTCCTCAATGGACAGGCCTTCTCTGGGGAGGTGCAGCTCAT  
 CCACTATAACCATGAGCTATATACGAATGTCACAGAAGCTGCAAAGAGTCCAAATGGATTGGT  
 GGTAGTTTCTATATTTATAAAAGTTTCTGATTTCATCAAACCCATTTCTTAATCGAATGCTCAA  
 CAGAGATACTATCACAAGAATAACATATAAAAATGATGCATATTTACTACAGGGGCTTAATAT  
 AGAGGAACATATCCAGAGACCTCTAGTTTCATCACTTACGATGGGTGATGACTATCCCACC  
 CTGCTATGAGACAGCAAGTTGGATCATAATGAACAAACCTGTCTATATAACCAGGATGCAGAT  
 GCATTCCCTTGCGCCTGCTCAGCCAGAACCAGCCATCTCAGATCTTCTGAGCATGAGTGACAA  
 CTTCAGGCCTGTCCAGCCACTCAACAACCGCTGCATCCGCACCAATATCAACTTCAGTTTACA  
 GGGGAAGGACTGTCCAAACAACCGAGCCAGAAGCTTCAGTATAGAGTAAATGAATGGCTCCT  
 CAAG**TAG**GGAACAAAGCCAAGAAGAATCCCACCTCAGTGAAATGCTACAACCTGTGAATTGACG  
 TAACCTAGAATGTCCCCCTTCTTGCTTCTCTCTCCTTCTTTCCCCCAAGCCTCATTCATTCTT  
 GGGATTGGCCCTTCTTTCATGAAAAGTGTCTGCGAAACCATGGCAGAGGAATACATCTCTCAC  
 ACATACTCACAACACACACACAAGCACTTGCACATACATACAAACACATGCAAACATACCTA  
 CACACACACACTCTCTTACAACCTCCATCATGGGAAGTCAAGTTTCAGAAACAAAAGTCTCAT  
 TCATAAGAGGTCTTAGAAGAAAATAACCAGTTAACCTGATTTCAATTTTGATACCGTTTTCT  
 GAACTAATAAATCTACCCAATGAGACTTTTTAGCCTTTGTACATACAAAATTCTTCCAAAAGA  
 GAGAGGAGAAAATACAGCTCTGATGGCATCAAACGGACTTTGCATCAAGTAATTTAGATAGT  
 GTCCTAGGATCCTTTGAGGGTGTGGTAGCAGGTGAGCAGGACAAAGTTGACCAAGGACACTT  
 ATTTCTAGATTATGATTCTTCTGTTTACTCAACAATTTACAAAGAAAAAAGGACAGACATTG  
 AAGAGCTACACATTGTATATATATCACCACAGACTATAAGGAAATGGAATTATTTCCCTCTTT  
 GTCACATATCTGTAGTAGGATTTGCCAAGATCAGAAATGATCCATTTGCTGTTTCTTGTTTTT  
 CAAAGGTCATACATTGTGTTTGGTTATTGTTACCAGCTCAATAAATGTGTTTAAACGAGTTAAT  
 TTCATTTTTCTGGCTTTGGTCTGTTCTCCTTCCCTACAGGCTAAGCCCTGGCTCCATGCAACT  
 GCATTCTTTGATTTCACTTGTTCCCTTCATCTACATGTTTTGTTTCAATTTGCAGCCAGTTTTTAC  
 TGAGTTTGTGGCAATCAGGAATGCATTTGCTAAGCAAGTATGACTTTAATTCACCTCCATGGC  
 TCAATCATTCACATGAGGTGAGCTTCAGCCTGAGATAGCAGGCGACAGACTTCTTGCGTTTCA  
 AAAGTGGCATGCCCCCTGTGATGCTCCCGTGAAGGAATGCACTTTGCCTTGTAAGTTCCCTGG  
 GAAAGGGGTATGTTTTCTCTCCAGGTGCAGCCAGATCTCACAAGTACAAAACGAATGCCTTT  
 CTTTTCTTGTTTATAATGGTCACTCACTGTGTTTGGTTACTGTCAAGAAATCAATAAATGTGT  
 TTAACAAGTTA

**FIGURE 34**

MEIVWEVLFLLQANFIVCISAQQNSPKIHEGWWAYKEVVQGSFVPVPSFWGLVNSAWNLC SVG  
KRQSPVNIETSHMIFDPFLTPLRINTGGRKVS GMTYNTGRHVSRLRDKEHLVNI SGGPMTYSH  
RLEEIRLHFGSEDSQGSEHLLNGQAFSGEVQLIHYNHELYTNVTEAAKSPNGLVVVSIFIKVS  
DSSNPFLNRMLNRDTITRITYKNDAYLLQGLNIEELYPETSSFITYDGSMTIPPCYETASWII  
MNKPVYITRMQMHSRLRLLSQNQPSQIFLSMSDNFRPVQPLNNRCIRTNINFSLQKDCPNRA  
QKLQYRVNEWLLK

### Important features:

Signal peptide:

amino acids 1-20

Eukaryotic-type carbonic anhydrases proteins.

amino acids 126-162, 220-269, 43-91

N-glycosylation sites.

amino acids 116-119, 168-171, 302-305

GTCTGGAAACCCCTCAGGCCACCCTCGGGAGTCTCTGGGGTCCAGAGGGGTGTCCTGTACCCCTTGCAC  
ACAGGACCCTCACTCTGCAGGGATAAGCCAGCTGCGCCTGCAGCCTAGGGTGCCAAGGAGGCTGCTGA  
TTGTGGCCACAGCCTCATCTGAACGCCAGGAGACCAGGATACCGAGGCACCGGATCCCCTCTCTGTG  
CCCTGGGGAGCCCCAGTGCTGCCAGTCACCCCAGGGCTGAGGTCTGCGTCCCTAGTGGTGCAAGGCC  
TGGTAGGACCACGGGGCAGGGAATGTGAGCGCCATCCGAGCTCACGGTGTCTGAGTCGCGGCTTCGT  
GACTTTGGCAGGGGCCCTCCGGACCAGTGACCCAGTCAAACCCAGAGGGTCTTGGGCGGCAGCGACGA  
AGGAGGTATTAGGCTCCAGGCCAGGTGGGGCCGGACGCCCCCAGCCATCCACCATGGTGGTGGCACA  
CCCCACCGCCACTGCCACCACCACGCCCACTGCCACTGTACAGGCCACC GTTGTGATGACCACGGCCA  
CCATGGACCTGCGGGACTGGCTGTTCTCTGTACGGGCTCATCGCCTTCCTGACGGAGGTTCATCGA  
AGCACACCTGCCCCCTCGGTGTGCCGTGCGACAACGGCTTCATCTACTGCAACGACCGGGGACTCA  
ATCCATCCCCGCAGATATCCCTGATGACGCCACCACCTCTACCTGCAGAACACCAGATCAACAACG  
CCGCATCCCCCAGGACCTCAAGACCAAGGTCAACGTGCAGGTCACTTACCTATACGAAATGACCTG  
GATGAGTTCCCCATCAACCTGCCCCGCTCCCTCCGGGAGCTGCACCTGCAGGACAACAATGTGCGCAC  
CATTGCCAGGGACTCGCTGGCCCCGATCCCGCTGCTGGAGAAGCTGCACCTGGATGACAACCTCCGTGT  
CCACCGTCAGCATTGAGGAGGACGCCTTCGCCGACAGCAAACAGCTCAAGCTGCTCTTCCTGAGCCGG  
AACCACCTGAGCAGCATCCCCTCGGGGCTGCCGCACACGCTGGAGGAGCTGCGGCTGGATGACAACCG  
CATCTCCACCATCCCGCTGCATGCCTTCAAGGGCCTCAACAGCCTGCGGCGCCTGGTGCTGGACGGTA  
ACCTGCTGGCCAACCAGCGCATCGCCGACGACACCTTCAGCCGCCCTACAGAACCTCACAGAGCTCTCG  
CTGGTGCGCAATTGCTGGCCGCGCCACCCCTCAACCTGCCAGCGCCACCTGCAGAAGCTCTACCT  
GCAGGACAATGCCATCAGCCACATCCCCTACAACACGCTGGCCAAGATGCGTGAGCTGGAGCGGCTGG  
ACCTGTCCAACAACAACCTGACCACGCTGCCCCGCGGCCTGTTGACGACCTGGGGAACCTGGCCAG  
CTGCTGCTCAGGAACAACCCCTTGGTTTTGTGGCTGCAACCTCATGTGGCTGCGGGACTGGGTGAAGGC  
ACGGCGCGCGTGCTCAACGTGCGGGCCCTCATGTGCCAGGGCCCTGAGAAGGTCCGGGCATGGCCA  
TCAAGGACATTACCAGCGAGATGGACGAGTGTTTTGAGACGGGGCCGAGGGCGGCTGGCCATGGC  
GCTGCCAAGACCACGGCCAGCAACCACGCCTCTGCCACCACGCCCCAGGGTTCCCTGTTTACCCTCAA  
GGCCAAAAGGCCAGGGCTGCGCCTCCCCGACTCCAACATTGACTACCCCATGGCCACGGGTGATGGCG  
CCAAGACCCTGGCCATCCACGTGAAGGCCCTGACGGCAGACTCCATCCGCATCACGTGGAAGGCCACG  
CTCCCCGCCTCCTCTTTCCGGCTCAGTTGGCTGCGCCTGGGCCACAGCCAGCCGTGGGCTCCATCAC  
GGAGACCTTGGTGACAGGGGGACAAGACAGAGTACCTGCTGACAGCCCTGGAGCCCAAGTCCACCTACA  
TCATCTGCATGGTCACCATGGAGACCAGCAATGCCTATGTAGCTGATGAGACACCCGTGTGTGCCAAG  
GCAGAGACAGCCGACAGCTATGGCCCTACCACCACACTCAACCAGGAGCAGAACGCTGGCCCCATGGC  
GAGCCTGCCCCCTGGCGGGCATCATCGGCGGGGCAGTGGCTCTGGTCTTCCTCTTCCTGGTCTGGGGG  
CCATCTGCTGGTACGTGCACCAGGCTGGCGAGCTGCTGACCCGGGAGAGGGCCTACAACCGGGGCAGC  
AGGAAAAGGATGACTATATGGAGTCAGGGACCAAGAAGGATAACTCCATCCTGGAAATCCGCGGCCCT  
TGGGCTGCAGATGCTGCCATCAACCCGTACCCGCCCCAAGAGGAGTACGTGGTCCACACTATCTTCC  
CCTCCAACGGCAGCAGCCTCTGCAAGGCCACACACACCACTTGGCTACCGCACCGGGGCTACCGG  
GACGGCGGCATCCCCGACATAGACTACTCCTACACATGATGCCCGCCACCAGGGCTGCCCGCCTCA  
GCCCCAGCTGCCCTGGCGTGGCCATGTGGCTTTGCCAGCCTGCTGCAATCCAAGAGAGCAAGGAAGA  
GAAATTCCATGGGTGACTTTCTCCGCAGAAAGCAAAGTTTGGGGAGGGCTGACGATTTTGTAGAACA  
CAACAGTGACAATTTTTTTTAAAGAATAGAAGGCAGGAGGGGGAATTGACATTGTTGAAGACATAA  
TTTATACCAAGTTATGCCAGTTGGGGAGGGAAGGACTAAAAATAATATTGCAGGCAGGGCTGGGTTGG  
GTTTTTTTTTTTCCCCCTGAACTGGAAGGATACTACCTGTACAACATCTGTGGACACCTCATGCTCT  
GTTCAAGGCCATCACAAGGAACCGCCAGGGAGAAGCAGCCGGCTCTCAAAGCTCCACGCAGCTCTC  
CCGCCACTGGCCACTCGCTGGCGACCCGATGGAAGGTTTTTCAGGCTCCTCACAAGGAGAGAGGGAAG  
AAAAGATCTTTTGCCCTGGAGATATGGTCCTGAAATCTCTCCCCTGGCTTATTCCATACCATTTCCCT  
TGCAGATTGCGAGAAACATGGCATCTTCACTGCATTCTTTGAACAATCATGTAGTCGATTAAAAAA  
AAAACAAACTTTTTTTCTTAGGCTGAAGCCCTCTTCAGTTCCATGCACCACGCTCCGTAGAAGCCCC  
GGCGGAAGCCGTAGCTTTCCCTGCCACCTGGAGGTGCATCTGTCTGCCTGTCTATCCCTGTGCGGGTG  
TCTCTAAGTACAGATGGGTAGATAGAGCCACATGCACGCTCCTTACCGTTCTTCTTGGGTGAGTTCTT  
ACCATTTCTGAAACAATAGAATTGTGAAAGTGTTAAAAA

**FIGURE 36**

MVVAHPTATATTTPTATVTATVVMTTATMDLRDWLFLCYGLIAFLTEVIDSTTCPSVCRCDNG  
 FIYCNDRGLTSIPADIPDDATTLYLQNNQINNAGIPQDLKTKVNVQVIYLYENDLDEFPINLP  
 RSLRELHLQDNNVRTIARDSLARIPLLEKLHLLDDNSVSTVSI EEDAFADSKQLKLLFLSRNHL  
 SSIPSGLPHTLEELRLDDNRISTIPLHAFKGLNSLRRLVLDGNLLANQRIADDTFSRLQNLTE  
 LSLVRNSLAAPPLNLP SAHLQKLYLQDNAISHIPYNTLAKMRELERLDLSNNNLTTLP RGLFD  
 DLGNLAQLLLRNNPWFCGCNLMWLRDWVKARAAVVNV RGLMCQGPEKVRGMAIKDITSEMDEC  
 FETGPQGGVANAAAKTTASNHASATTPQGS LFTLKAKRPGLRLPDSNIDYPMATGDGAKTLAI  
 HVKALTADSIRITWKATLPASSFRLSWLRLGHSPAVGSITETLVQGDKTEYLLTALEPKSTYI  
 ICMVTMETSNA YVADETPVCAKAETADSYGPTTTLNQEQNAGPMASLPLAGIIGGAVALVFLF  
 LVLGAICWYVHQAGELLTRERAYNRGSRKKDDYMESGTTKDNSILEIRGPGLQMLPINPYRAK  
 EEYVVHTIFPSNGSSLCKATH TIGYGTTRGYRDGGIPDIDYSYT

**Important features of the protein:****Transmembrane domain:**

amino acids 552-573

**N-glycosylation sites.**

amino acids 249-252, 305-308, 642-645

**Leucine zipper pattern.**

amino acids 182-203, 299-320

**Phospholipase A2 aspartic acid active site.**

amino acids 57-67



[illegible]

## FIGURE 38

MAEPGSHSHLSARVRRRTERRIPRLWRLLLWAGTAFQVQTQGTGPELHACKESEYHYEYTACDS  
TGSRWVRVAVPHTPGLCTSLSDPVKGTECSFSCNAGEFLDMKDQSCKPCAEGRYSLGTGIRFDE  
WDELPHGFASLSANMELDDSAESTGNCTSSKWVPRGDYIASNTDECTATLMYAVNLKQSGTV  
NFEYYPDPSSIIFFEFFVQNDQCQPNADDSRWMTTEKGWEFHSVELNRGNVLYWRTTAFSVW  
TKVPKPVLVRNIAITGVAYTSECFPCPGTYADKQGSSFCKLCPANSYSNKGGETSCHQCDPDK  
YSEKSSSSCNVRPACTDKDYFYHTTACDANGETQLMYKWAKPKICSEDLEGAVKLPAASGVKTH  
CPPCNPGGFKTNNSTCQPCPYGSYNGSDCTRCPAGTEPAVGFEYKWWNTLPTNMETTTLVLSGI  
NFEYKGMTGWEVAGDHIYTAAGASDNDFMILTLVVPGRPPQSVMAADTENKEVARITFVFETL  
CSVNCELYFMVGVNSRTNTPVETWKGSKGKQSYTYIIIEENTTTSTFWAFORTTFHEASRKYTN  
DVAKIYSINVTNVMNGVASYCRPCALEASDVGSSTSCPAGYYIDRDSGTCHSCPPNTILKAH  
QPYGVQACVPCGPGTKNNKIHSCLYNDCTFSRNTPTRTFNYNFSALANTVTLAGGPSFTSKGL  
KYFHHFTLSLCGNQGRKMSVCTDNVTDLRIPEGESGFSKSIYAVVCQAVIIPPEVTGYKAGVS  
SQPVSLADRLIGVTTDMTLDGITSPAELFHLESLGIPDVIFFYRSNDVTSQSCSSGRSTTIRVR  
CSPQKTVPGLSLLLPGTCSDGTCDCGNFHFLEWESAAACPLCSVADYHAIVSSCVAGIQXTTYVX  
REPKLCSGGISLPEQRTVICTIDFWLKVGISAGTCTAILLTVLTCYFWKKNQKLEYKYSKLV  
MNATLKDCLPAADSCAIMEGEDVEDDLIFTSKKSLFGKIKSFTSKRTPDGFDSVPLKTS  
SGGPDMDL

Important features of the protein:

N-glycosylation sites:

amino acids 153-156, 390-393, 391-394, 404-407, 544-547, 576-579,  
672-675, 717-720, 947-950

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

amino acids 15-18, 563-566, 709-712

Casein kinase II phosphorylation sites:

amino acids 42-45, 59-62, 81-84, 146-149, 168-171, 282-285, 331-334, 340-343, 431-434, 449-452, 465-468, 523-526, 557-560, 761-764, 780-783, 835-838, 860-863, 893-896, 949-952

**Tyrosine kinase phosphorylation sites:**

amino acids 50-56, 109-116

N-myristoylation sites:

amino acids 77-82, 88-93, 152-157, 268-273, 288-293, 320-325,  
400-405, 405-410, 414-419, 463-468, 599-604, 616-621, 634-639,  
644-649, 839-844, 874-879, 912-917, 916-921

Amidation site:

amino acids 707-710

Cell attachment sequence:

amino acids 162-164

**FIGURE 39**

GGGAAGGGGTTCTGGGCTGCCGCAGGCACACAGGCCAGAGCTTCGTGGATACCTGCAGGGCCC  
AAAGGTCCCTCCCTGTTTTGAAGAGTGAGTGATGGCTATGAGGTAGCGGCCAGGCTGATCACC  
CCTGCGTTGGCTGGAGGCAGAATTCTGTAAATCCTCGCCAAGTCTTTCTCCAGGGCCACTGGTT  
AGCTCATCTCAGCCTCCTCTGGGAGCATCAACACCAACATGGCACAGGGGACTGCAGTGGTGT  
GCTTTGGACCTGTGTACCCACCCAAGGCTAAAGGCAGAGCCAGGTGACTTTGCGGGGGTCTCT  
TCTCTAGGATTATCTGTACTTCCCCTCTGTCTCTTTTACTACGGGAGATCGAGCTAGCTATA  
ACCCACCTTCTTTCATGAGAACCACACTAAATTGCAAAAATTATCCCAGTGCTGGAGGAGGGC  
AGCAGGTTGAGATTATGTTGGCAGGAAGAATGTTGGCATTGATTGGCACGCAGGGGACGAGAG  
CTGCTTTGTGCTTTAAAGGAGCCAAGTTACACCCTGTTTAAACCCTGCCTTCAAAGGGACGACT  
CTGTAAGATTCTCTGCTACTTATTCAAGTTGACACGATGCCCTTCCACTCCACCTGAGGTCC  
CGCCTTCCCTCTGCCATAAGGAGTTTGATTCTACAAAAGAAACCAAAACATCAGAAATACATCC  
AGCATGGCTGGAGAGCTCCGACCAGCCAGCCTGGTGGTCCTGCCAGGTCCCTTGCTCCAGCT  
TTTGAAAGATTCTGCCAGGTCAACACTGGTCCCTCTACCCCTGCTGGGCCAGAGTGAGCCAGAA  
AAGTGGATGCTGCCCCCTCAAGGTGCTATCTCAGAGACCAGGATGGGCCATCCCCAGTTCTGG  
AAATACGAGTTCGGTGCCTGCACCGGTAGCCTGGCTTCGCTGGAGCAGTACTCGGAGCAGCTG  
AAGGACATGGTGGCCTTCTTCCCTGGGCTGCAGCTTCTCCCTGGAGGAGGCCTTGAGAAAGCG  
GGGCTCCCCAGAAGAGACCCAGCAGGTACAGCCAGGCGGGTGCATACAAGACAACAGTGCCT  
TGTGTTACCCATGCTGGCTTCTGCTGCCCTCTGGTGGTCACGATGAGGCCCATTTCCAAGGAC  
AAGCTGGAAGGGCTGGTGCGGGCCTGCTGCTCCCTCGGAGGTGAGCAGGGGCAACCTGTTTAC  
ATGGGCGACCCAGAACTGTTGGGAATCAAAGAGCTTTCCAAACCTGCCTACGGGGATGCCATG  
GTGTGTCCCCCAGGGGAGGTTCCAGTGTTCTGGCCTTCTCCGCTGACCAGTCTCGGAGCTGTC  
AGCAGCTGTGAGACCCCACTGGCTTTTGCCAGCATCCCAGGCTGCACAGTTATGACTGACCTG  
AAGGATGCAAAGGCTCCACCTGGTGTCTACCCCCAGAGAGAATTCCAGAGGTCCATCACATT  
TCCCAAGATCCTCTGCACTACAGCATCGCGTCAGTCTCTGCTTCTCAGAAGATCAGAGAATA  
GAGTCTATGATCGGCATAGACCCAGGGAACCGGGGGATTGGGCACCTGCTCTGTAAAGATGAG  
CTGCTGAAGGCCTCTCTCTCGCTGTCCCATGCCCGCTCAGTGCTCATCACCCTGGGTTCCCC  
ACACATTTCAATCATGAGCCTCCAGAAGAGACAGATGGCCCACCAGGAGCTGTTGCTCTGGTT  
GCCTTCCCTGCAGGCCTTGGAAGAGGAGGTCGCCATAATCGTTGACCAGAGAGCCTGGAACCTG  
CACCAGAAGATTGTTGAAGATGCTGTTGAGCAAGGTGTTCTGAAGACGCAGATCCCAGATTA  
ACTTACCAAGGTGGATCAGTGGAAGCTGCTCAGGCATTCTGTGCAAAAATGGGGACCCGCAG  
ACACCTAGATTTGACCACCTGGTGGCCATAGAGCGTGCCGGAAGAGCTGCTGATGGCAATTAC  
TACAATGCAAGGAAGATGAACATCAAGCACTTGGTTGACCCCATTGACGATCTTTTTCTTGCT  
CGAAGAAGATTCTGGAATCTCATCAACTGGAGTCGGTGATGGAGGCAACGAGCTTGGGATG  
GGTAAAGTCAAGGAGGCTGTGAGGAGGCACATACGGCACGGGGATGTCATCGCCTGCGACGTG  
GAGGCTGACTTTGCCGTCATTGCTGGTGTCTTAAGTGGGGAGGCTATGCCCTGGCCTGCGCA  
CTCTACATCCTGTACTCATGTGCTGTCCACAGTCAGTACCTGAGGAAAGCAGTCGGACCCCTCC  
AGGGCACCTGGAGATCAGGCCTGGACTCAGGCCCTCCCGTCGGTCATTAAGGAAGAAAAATG  
CTGGGCATCTTGGTGCAGCACAAAGTCCGGAGTGGCGTCTCGGGCATCGTGGGCATGGAGGTG  
GATGGGCTGCCCTTCCACAACACCCACGCCGAGATGATCCAGAAGCTGGTGGACGTCACCACG  
GCACAGGTGTAACCGTCCATGTTCCGTGTGAGCAGAGTCCCTACCAACGGGCAGGTCTGCATC  
CGGGGAGAATGCAGCTGCTTCTGGCGACAATCCTGCTAGTAAACACTGGTCTTCCGTGAGCAA  
CGAACACTCGCCTGGCCTGGGAACTGCATGCCCACTTTCTGGGAGGGGTTAGTGCAGGTGCC  
GTGGACAAAGGACAACATTTCTCTGGGGCTTTTTAACTTTTATTCTTAAGACTCTAAAGGCGT  
TGATTTCAACCCTCCTTCACTCTGGCTTCTTCAGGCAACCCACGTGGTCTCCTATGAGAATCT  
TCTCGACAGTTACTTATGGGGACACTTGTGAACAATTAAGTCCAGGGGCAGAGCATGAGAACA  
AACATTTCCAGGCCATGTAGGATAGGATACTCCAGACTCCAGTCATCCTCCCCCATCCATGGT  
TTCTGTTACTCATGGTTTCAGTTACTCATAGCCAAGTGCAGACCGAAAATACTAAATGAAAAA  
TTTCAGAAATAAACAACTCTTAAGTTTTAAAAAAA

**FIGURE 40**

MPFTLHLRSRLPSAIRSLILQKKPNIRNTSSMAGELRPASLVVLPRSLAPAFERFCQVNTGPL  
PLLGQSEPEKWMLPPQGAISETRMGHPQFWKYEFGACTGSLASLEQYSEQLKDMVAFFLGCSF  
SLEEALAKAGLPRRDPAGHSQAGAYKTTVPCVTHAGFCCPLVVTMRPIPKDKLEGLVRACCSL  
GGEQQQPVHMGDPPELLGIKELSKPAYGDAMVCPGGEVPVFWPSPLTSLGAVSSCETPLAFASI  
PGCTVMTDLKDAKAPPGCLTPERIPEVHHISQDPLHYSIASVSASQKIRELESMIGIDPGNRG  
IGHLLCKDELLKASLSLSHARSVLITTFPTFNHEPPEETDGPPGAVALVAFLQALEKEVAI  
IVDQRAWNLHQKIVEDAVEQGVLTQIPILTYQGGSVAAQAFCKNGDPQTFRFDHLVAIER  
AGRAADGNYYNARKMNIKHLVDPIDDLFLAAKKIPGISSTGVGDGGNELGMGKVKEAVRRHIR  
HGDVIACDVEADFAVIAGVSNWGGYALACALYILYSCAVHSQYLRKAVGPSRAPGDQAWTQAL  
PSVIKEEKMLGILVQHKVRSGVSGIVGMEVDGLPFHNTHAEMIQKLVDVTTAQV

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 358-378, 517-539

**N-glycosylation site.**

amino acids 28-32

**Tyrosine kinase phosphorylation site.**

amino acids 444-452

**N-myristoylation site.**

amino acids 98-104, 102-108, 123-129, 149-155, 181-187, 190-196,  
238-244, 308-314, 399-405, 413-419, 448-454, 477-483, 482-488,  
487-493

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 233-244, 531-542

**FIGURE 41**

CTTTCCTGTTTATACCGCAGCCCTTTTCTTCTTTGAGTTAGTAAAGATTTATTCTGTAACCTG  
ACACTCATCTGGCCCTTTGCAGTTTGCCAGCCATATTCCCATGTGATTTCCCAGTGGATCCAG  
GCCCCATCCGGCTGGCAGGAGGGGGCTCTGACGTACAGGTTGGAAATCAGAAGTCTGTGAGA  
GCGCGGGAGTGCATGGCAGCTCTGGGTCCCAGACCTGGCCCCGACCCCTCTGCTTCACCTCCAG  
CTCTGCTGCTCCTCTACTCTTGGGTCGAGATCCCTTTGGAGCCACAGCGAGGAACCCCTGTGGT  
CCTCAGGCAGGTGTACCTTGAGTCAGCCAGGAGCCCTCTTTTCCTGTGTCAAAGCCTGCCCTC  
GGGCTCTGCTCACCTCTGGTGACCCCTCCAAGATGCCCCTGCCCTCAGTTTCCCCTCATGATCT  
GGCCTCTGCCCCCTTCTCTAGCCACAGCCTCTAGTACACTTTAGCAATACCACCAGACTAGTT  
AGAGTTCCCCACTACCAAGCAAGACATGCAGTTTTCATGCCTCTGTGCCTTCGCTCATGCTGT  
TTCTTCCGACTGGAATGCCTTCCCCTGCTCCTCCTGCCTTGTCTGCCTGGCAAGTTCATCTCT  
CACGATCCCCCTCAAAGGCCCCCTCCTCCAGGAAGGCAACCCCTGTGCCCTCCCCTCCAGGCT  
ACCTCTGCACTTTGTCAATGCTTCTCTTGTGGCACTTATCACACTGTATTTTACTTGTTTACA  
TGTTTGTCTCCCCTTCTAGACTGTGAATCCTTAAGGGCATGGACTGTATCTTATGCATCTCTG  
TATTTCTGCGCCTAGCACGGTGCCCTAGCACACAGTAGGCGCTCAATAAATGTTGAATGAATGA  
ATGATTT

**FIGURE 42**

MQFHASVPSLMLFLPTGMPSAPPALSAWQVHLSRSPQRPPPPGRQPLCPSPPGYLCTLSMLL  
LWHLSHCILLVYMFVSPSRL

Important features of the protein:

Signal peptide:

amino acids 1-22

Microbodies C-terminal targeting signal.

amino acids 81-83

**FIGURE 43**

GTTTCCAACAAGGATGATATGAAGACTTCCCTGAAGAAAGTTGTGAAGGGACCTCCTACGAGA  
**TG**ATGATGCAGTGTGTGTCCCGCATGTTGGCCCACCCCTGCATGTCATCTCAATGCGCTGCA  
 TGGTCCAGTTTGTGGGACGGGAGGCCAAGTACAGTGGTGTGCTGAGCTCCATTGGGAAGATTT  
 TCAAAGAGGAAGGGCTGCTGGGATTCTTCGTTGGATTAATCCCTCACCTCCTGGGCGATGTGG  
 TTTTCTTGTGGGGCTGTAACTTGCTGGCCCACTTCATCAATGCCTACCTGGTGGATGACAGCT  
 TCAGCCAGGCCCTGGCCATCCGGAGCTATACCAAGTTCGTGATGGGGATTGCAGTGAGCATGC  
 TGACCTACCCCTTCCTGCTAGTTGGCGACCTCATGGCTGTGAACAACCTGCGGGCTGCAAGCTG  
 GGCTCCCCCTTACTCCCCAGTGTTCAAATCCTGGATTCACTGCTGGAAGTACCTGAGTGTGC  
 AGGGCCAGCTCTTCCGAGGCTCCAGCCTGCTTTTCCGCCGGGTGTCATCAGGATCATGCTTTG  
 CCCTGGAG**TAA**CCTGAATCATCTAAAAAACACGGTCTCAACCTGGCCACTGTGGGTGAGGCCT  
 GACCACCTTGGGACACCTGCAAGACGACTCCAACCCAACAACAACCAGATGTGCTCCAGCCCA  
 GCCGGGCTTCAGTTCCATATTTGCCATGTGTCTGTCCAGATGTGGGGTTGAGCGGGGGTGGGG  
 CTGCACCCAGTGGATTGGGTCAACCGGCAGACCTAGGGAAGGTGAGGCGAGGTGGGGAGTTGG  
 CAGAATCCCCATACCTCGCAGATTTGCTGAGTCTGTCTTGTGCAGAGGGCCAGAGAATGGCTT  
 ATGGGGGGCCCAGGTTGGATGGGGAAAGGCTAATGGGGTCAGACCCACCCCGTCTACCCCTCC  
 AGTCAGCCCAGCGCCCATCCTGCAGCTCAGCTGGGAGCATCATTCTCCTGCTTTGTACATAGG  
 GTGTGGTCCCCTGGCACGTGGCCACCATCATGTCTAGGCCTATGCTAGGAGGCCAAATGGCCAG  
 GCTCTGCCTGTGTTTTTCTCAACACTACTTTTCTGATATGAGGGCAGCACCTGCCTCTGAATG  
 GGAAATCATGCAACTACTCAGAATGTGTCCTCCTCATCTAATGCTCATCTGTTTAATGGTGAT  
 GCCTCGCGTACAGGATCTGGTTACCTGTGCAGTTGTGAATACCCAGAGGTTGGGCAGATCAGT  
 GTCTCTAGTCCTACCCAGTTTTTAAAGTTCATGGTAAGATTTGACCTCATCTCCCGCAAATAAA  
 TGTATTGGTGATTTGGAAAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MMMQCVSRMLAHPLHVISMRCMVQFVGREAKYSGVLSSIGKIFKEEGLLGFFVGLIPHLLGDV  
VFLWGCNLLAHFINAYLVDDSFSQLAIRSYTKFVMGIAVSMLTYPFLLVGDLMVNNCGLQA  
GLPPYSPVFKSWIHCWKYLSVQGQLFRGSSLLFRRVSSGSCFALE

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-18

#### **Transmembrane domains:**

amino acids 51-72, 97-114

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 160-163

#### **N-myristoylation sites.**

amino acids 34-39, 100-105, 123-128, 165-170



GCTCACTCTTTGGGTGCCACACTGCCTTTATGAGACTGTAAACACTCACTGGGAATGTCTGCAGCT  
TCACTCCTGAAGCCAGCGAGACCAGAACCCACCAGGAGGAACAACTCCAGACGCGCAG  
CCTTAAGAGCTGTAACACTCACCGCGAAGGTCTGCAGCTTCACTCCTGAGCCAGCCAGACCAC  
GAACCCACCAGAAGGAAGAACTCCAACACATCCGAACATCAGAAGGAGCAAACCTCGTGACA  
CGCCACCTTTAAGAACCGTGACACTCAACGCTAGGGTCCGCGGCTTCATTCTTGAAGTCAGTG  
AGACCAAGAACCCACCAATTCCGGACACGGCAAAGTAACATCCTAGACATGGCTTTAGAGATC  
CACATGTCAGACCCCATGTGCCTCATCGAGAACTTTAATGAGCAGCTGAAGGTTAATCAGGAA  
GCTTTGGAGATCCTGTCTGCCATTACGCAACCTGTAGTTGTGGTAGCGATTGTGGGCCTCTAT  
CGCACTGGCAAATCCTACCTGATGAACAAGCTGGCTGGGAAGAACAAAGGGCTTCTCTGTTGCA  
TCTACGGTGCAGTCTCACACCAAGGGAATTTGGATATGGTGTGTGCCTCATCCCAACTGGCCA  
AATCACACATTAGTTCTGTCTTGACACCGAGGGCCTGGGAGATGTAGAGAAGGCTGACAACAAG  
AATGATATCCAGATCTTTGCACTGGCACTCTTACTGAGCAGCACCTTTGTGTACAATACTGTG  
AACAAAATTGATCAGGGTGCTATCGACCTACTGCACAATGTGACAGAAGCTGACAGATCTGCTC  
AAGGCAAGAACTCACCTGACCTTGACAGGGTTGAAGATCCTGCTGACTCTGCGAGCTTCTTC  
CCAGACTTAGTGTGGACTIONTGTAGAGATTTCTGCTTAGGCCTGGAAATAGATGGGCAACTTGT  
ACACCAGATGAATACCTGGAGAATTCCTAAGGCCAAAGCAAGGTAGTGATCAAAGAGTTCAA  
AATTTCAATTTGCCCGTCTGTGTATACAGAAGTTCTTTCCAAAAAGAAATGCTTTATCTTT  
GACTTACCTGCTCACCAAAAAAAGCTTGCCCAACTTGAAACACTGCCTGATGATGAGCTAGAG  
CCTGAATTTGTGCAACAAGTGACAGAATTCGTGTTCTTACATCTTTAGCCATTCTATGACCAAG  
ACTCTTCCAGGTGGCATCATGGTCAATGGATCTCGTCTAAAGAACCTGGTGTGACCTATGTC  
AATGCCATCAGCAGTGGGGATCTGCCTTGCAATAGAGAATGCAGTCTTGCCCTTGGCTCAGAGA  
GAGAAGCTCAGCTGCAGTGCAAAAGGCCATTGCCACTATGACCAGCAAATGGGCCAGAAAGTG  
CAGCTGCCCATGGAAACCCTCCAGGAGCTGCTGGACCTGCACAGGACCAGTGAGAGGGAGGCC  
ATTGAAGTCTTCATGAAAACTCTTTCAAGGATGTAGACCAAAGTTTCCAGAAAGAATTGGAG  
ACTCTACTAGATGCAAAACAGAATGACATTTGTAACCGGAACCTGGAAGCATCTCGGATTAT  
TGCTCGGCTTTACTTTAAGGATATTTTTGGTCTCTAGAAGAAGCAGTGAAGCAGGGAATTTAT  
TCTAAGCCAGGAGGCCATAATCTCTTCATTCAGAAAACAGAAGAAGCTGAAGGCAAAGTACTAT  
CGGGAGCCTCGGAAAGGAATACAGGCTGAAGAAGTTCTGCAGAAATATTTAAAGTCCAAGGAG  
TCTGTGAGTCATGCAATATTACAGACTGACCAGGCTCTCACAGAGACGGAAAAAAGAAGAAA  
GAGGCACAAGTGAAAGCAGAAGCTGAAAAGGCTGAAGCGCAAAGGTTGGCGGCGATTCAAAGG  
CAGAACGAGCAAATGATGCAGGAGAGGGAGAGACTCCATCAGGAACAAGTGAGACAAATGGAG  
ATAGCCAAACAAAATTGGCTGGCAGAGCAACAGAAAATGCAGGAACAACAGATGCAGGAACAG  
GCTGCACAGCTCAGCACAAACATTCGAAGCTCAAAATAGAAGCCTTCTCAGTGAGCTCCAGCAC  
GCCCAGAGGGCTGTTAATAACGATGATCCATGTGTTTTACTCTTAAAGTGCTAAATATGGGAGT  
TTCCTTTTTTTTACTCTTTGTCACTGATGACACAACAGAAAAGAACTGTAGACCTTGGGACAA  
TCAACATTTAAATAAACTTTTATAATTATTAAA

**FIGURE 46**

MALEIHMSDPMCLIENFNEQLKVNQEALEILSAITQPVVVVAIVGLYRTGKSYLMNKLAKGNK  
 GFSVASTVQSHTKGIWIWCVPHPNWPNHTLVLLDTEGLGDVEKADNKNDIQIFALALLSSTF  
 VYNTVKNIDQGAIDLLHNVTELTDLLKARNSPDLDRVEDPADSASFFPDLVWTLRDFCLGLEI  
 DGQLVTPDEYLENSLRPKQGSQQRVQNFNLPRLCIQKFFPKKKCFIFDLPAHQKLAQLETLP  
 DDELEPEFVQQVTEFCSYIFSHSMTKTLPGGIMVNGSRLKNLVLTYVNAISSGDLPCIEHAVL  
 ALAQRENSAAVQKAIHYDQQMGQKVQLPMETLQELLDLHRTSERAIEVFMKNSFKDVDQSF  
 QKELETLLDAKQNDICKRNLEASSDYCSALLKDIFGPLEEAVKQGIYSKPGGHNLFIQKTEEL  
 KAKYYREPRKGIQAEVLQKYLKSKESVSHAILQTDQALTETEKKKKEAQVKAEAEKAEQRL  
 AAIQRQNEQMMQERERLHQEQVRQMEIAKQNWLAEQQKMQEQQMQEQAAQLSTTFQAQNRSL  
 SELQHAQRAVNNDPCVLL

**Important features of the protein:****Transmembrane domains:**

amino acids 31-49, 114-131

**N-glycosylation sites.**

amino acids 90-94, 144-148, 287-291, 563-567

**N-myristoylation sites.**

amino acids 45-51, 283-289

**Prenyl group binding site.**

amino acids 583-588

**ATP/GTP-binding site motif A (P-loop).**

amino acids 45-53

**FIGURE 47**

CACTCATTCAATCCAAAGGGTCTCTCAAGGCAATGGTAATGTGCAAGGAGGTGATACCTAAAT  
GAATGACCAAAGAACATGCTTCTGCTTTTGTGTGTCTCTACATTTTAGACATTTGTTTGT  
TCTCTTGGTAGCCTTTAAATTCCTTGAAGCCCAGGACCATGTCTCACTTACCTTTGTGTTTCC  
ACTAACTAGTCTACCTCCTGGAATTGGCAGATACTCAGTGAAAGCCTGTGAAATAAGTGATGT  
CTATTTCTAGCATATTATTCTGAGATTTAATGATAGATTTAGTGATTGAATGAGATTTCCATT  
TTCAAATACAGCAAAAGCATAACTATTTTCATTCATTCATATTCATTCAACTTCATTCTCAA  
ATTAGGTCCTGAGTTAACTAATAATTACCTTTGAAATGTGTGGGTATTGAGGCAATCAGGT  
GGTGACATTGAGCTCTCAGCCAGAGTTTGTCTTCTGGAATTGATTCAGTTCCATTGCATTGATT  
TTTGTCTCTCAGAAGCCAAGGTTTCCCATGAAAAATCATTCCCACTTGAATTGGGCTGTGATT  
TTGCTGCGTTTAAGTAAAGGAAGCCTCTTGGTTCTAGTTCTGCAAACCTACACACTGAACTGG  
GACAAGTTTTTGTTTAGAGTAATGGCTGGGAAAAGAGGAACCTTTCATTTTATTGAGAAGTCA  
AAAACAAAGGCCTCCAGCCACCTGGAGATGTTTTGTTGCAGACACCAGCCTGGCTCTGTCTT  
TATGCCTAACAATTGAGCATCCAGTCTTCTTTGTGCTGGGACCATTGCTCAGCTCTGCAAGGG  
GAAAAGAGGGGAGAAAGCCAGAGCTGCCAGGCTTCTTGCACTGGGGCCGGGGGAGGGTTCTTG  
GAAGCAGGTGCTCTCTGGCTTCTTGGTACGTGAGGCTCTCGGAGCTGCCTCTCCTCTGACCT  
CAGGTCCTCACCGAGTTTGCTCCAGGAGTATATTGAAAACATACCCAGTGCTCTCTCAAGCAC  
CCACTGCTTAGAGGGCCCAGATTTCTTTTCTTCTTTCCCTTGCGAGAGCTGGAGACTGCATCG  
GGCATCTGGTGTTTAAACTAAACAGGAAACTGACTAAAGGTCCACAGTGCTCATTGTGTAGA  
CTAGCTGCCCTCCG**ATG**GGTGCTCTGATTATCAGTGGTTCAGTGCAAGGGCCTGTCACTAAAC  
AGGCCTCACTTCCTCCTTGGGGGCTTTCCCATGGGAGGTGTGGCTTTTTACTCTACATGGAAA  
TGACTCTCTGCAGCCACAGAACACAGTCATTTTCTGAATTATCCAGTCTCTCATGCGCCCTG  
GATTCCTCCAGATGCCTTATATCTCTTGTGCAAAGTTGTCTAAAATTTGGTTCCAGCTTCCA  
AGCCTTGCCTTTTGGCCTTCTTGAAGTATTTTTGTTGATGAGTCGTCTGTCAATTATTCTCTA  
AAATGATTTGCTTTTTGTTTCTTTCATTCCCTATTTCCACCCACATATACACACATGCTTCT**T**  
**AA**CTTAGGGGATTACATGCCAATAAATCTATTGTTGAAAATGCACTAATACTATCGCAAAGAC  
GAAAATTCACAGGCTGAACCGTTGTAAGTCCATATGCTCCTCAACTTACATGTGTGATGGAGT  
TATGCCCAAATAAGTCCATCGTCAAGTTGAAAAATCAAAATCAAGCCATCTTAGGTTGAGGAC  
CATTTGTTTGTACCTCCAAAGATGTCATATCTTTAAACATACTCCCTAGCTTTTCTTTTACT  
TTTTATTTTGAAGTAATTATAGAATCACAGAAAGTTGCAAAAAA

## **FIGURE 48**

MGALIIISGSSAGPVTQASLPPWGLSHGRCGFLLYMEMTLCSHRTQSFSELSQSLMRPGFLQM  
PYISCAKLSKIWFPAKPCLLAFLEVFLMSRLSLFSKMICFLFLSFLFPPHIYTHAS

**Important features of the protein:**

**Signal peptide:**

amino acids 1-41

**Transmembrane domain:**

amino acids 88-107

**Casein kinase II phosphorylation site.**

amino acids 47-50

**N-myristoylation site.**

amino acids 24-29



**FIGURE 50**

MGPVVPSLGLLEGAPTRMVAAAVLQASRNPASTGQGPRCRES PGLLVVSGGKTNSLGQGRPPT  
PRPLENGHGGRSLGPGPLDWVEMPDHQRHPSTAPPTDLTSHLSRISLAGVEPSLVQAALGQLV  
RLSCSDDTAPESQAAWQKDGQPISSDRHRLQFDGSLIIHPLQAEDAGTYSCGSTRPGRDSQKI  
QLRIIGGDMAVLSEAELSRFPQPRDPAQDFGQAGAAGPLGAIPSSHPPQANRLRLDQNPQPRVV  
DASPGQRIRMTCAEGFPPPAIEWQRDQGPVSSPRHQLQPDGSLVISRVAVEDGGFYTCVAFN  
GQDRDQRWVQLRVLGELTISGLPPTVTVPEDGTARLLCVVAGESVNIRWSRNLFPVQADGHRV  
HQSPDGTLLIYNLRARDEGSYMC SAYQGSQAVSRSTEVKVVSPAPTAQPRDPGRDCVDQPELA  
NCDLILQAQLCGNEYYS SFCCASCSRFQPHAQPIWQ

**Important features of the protein:**

**Signal peptide:**

amino acids 1-16

**Tyrosine kinase phosphorylation site.**

amino acids 392-400

**N-myristoylation sites.**

amino acids 9-15, 50-56, 112-118, 146-152, 173-179, 195-201,  
220-226, 229-235, 280-286, 306-312, 336-342, 397-403

**Myelin P0 protein.**

amino acids 153-182

FIGURE 51

CAGGCAGAAGCGAACAAAGACCCAGCAAGAGAAGGCAGAGGGCTAAGACCCATCCCCTATCTGC  
TCTCCTGAAATAATTCTGGAGTCA**ATG**CCTGAAATGCCAGAGGACATGGAGCAGGAGGAAGTTA  
ACATCCCTAATAGGAGGGTTCTGGTTACTGGTGCCACTGGGCTTCTTGGCAGAGCTGTACACA  
AAGAATTTTCAGCAGAATAATTGGCATGCAGTTGGCTGTGGTTTTCAGAAGAGCAAGACCAAAAT  
TTGAACAGGTTAATCTGTTGGATTCTAATGCAGTTCATCACATCATTTCATGATTTTCAGCCCC  
ATGTTATAGTACATTGTGCAGCAGAGAGAAGACCAGATGTTGTAGAAAATCAGCCAGATGCTG  
CCTCTCAACTTAATGTGGATGCTTCTGGGAATTTAGCAAAGGAAGCAGCTGCTGTTGGAGCAT  
TTCTCATCTACATTAGCTCAGATTATGTATTTGATGGAACAAATCCACCTTACAGAGAGGAAG  
ACATACCAGCTCCCCTAAATTTGTATGGCAAAACAAAATTAGATGGAGAAAAGGCTGTCCTGG  
AGAACAATCTAGGAGCTGCTGTTTTGAGGATTCCCTATTCTGTATGGGGAAGTTGAAAAGCTCG  
AAGAAAGTGCTGTGACTGTTATGTTTTGATAAAGTGCAGTTCAGCAACAAGTCAGCAAACATGG  
ATCACTGGCAGCAGAGGTTCCCCACACATGTCAAAGATGTGGCCACTGTGTGCCGGCAGCTAG  
CAGAGAAGAGAATGCTGGATCCATCAATTAAGGGAACCTTTCAGTGGTCTGGCAATGAACAGA  
TGAATAAGTATGAAATGGCATGTGCAATTGCAGATGCCTTCAACCTCCCCAGCAGTCACTTAA  
GACCTATTACTGACAGCCCTGTCCTAGGAGCACACGTCGAGAAAATGCTCAGCTTGACTGCT  
CCAAATTGGAGACCTTGGGCATTGGCCAACGAACACCATTTCGAATTGGAATCAAAGAATCAC  
TTTGGCCTTTTCTCATTGACAAGAGATGGAGACAAACGGTCTTTTCAT**TAG**TTTTATTTGTGTTG  
GGTTCCTTTTTTTTTTTTAAATGAAAAGTATAGTATGTGGCACTTTTTTAAAGAACAAAGGAAATA  
GTTTTGTATGAGTACTTTAATTGTGACTCTTAGGATCTTTCAGGTAAATGATGCTCTTGCCT  
AGTGAAATTGTCTAAAGAACTAAAGGGCAGTCATGCCCTGTTTGCAGTAATTTTTCTTTTTTA  
TCATTTTGTGTTGTCCTGGCTAACTTGGAGTTTGTAGTATAGTAAATTATGATCCTTAAATATT  
TGAGAGTCAGGATGAAGCAGATCTGCTGTAGACTTTTCAGATGAAATTGTTTCATTCTCGTAAC  
CTCCATATTTTCAGGATTTTTGAAGCTGTTGACCTTTTCATGTTGATTATTTTAAATTGTGTG  
AAATAGTATAAAAATCATTGGTGTTTATTATTTGCTTTGCCTGAGCTCAGATCAAATGTTTG  
AAGAAAGGAACCTTTATTTTTTGCAAGTTACGTACAGTTTTTATGCTTGAGATATTTCAACATGT  
TATGTATATTGGAACCTTCTACAGCTTGATGCCTCCTGCTTTTATAGCAGTTTATGGGGAGCAC  
TTGAAAGAGCGTGTGTACATGTATTTTTTTTTCTAGGCAAACATTGAATGCAAACGTGTATTTT  
TTTAATATAAATATATAACTGTCCTTTTCATCCCATGTTGCCGCTAAGTGATATTTTCATATGT  
GTGGTTATACTCATAATAATGGGCCTTGTAAGTCTTTTCACCATTTCATGAATAATAATAAATA  
TGTACTGCTGGCATGTAATGCTTAGTTTTCTTGATTTTACTTCTTTTTTTTAAATGTAAGGACC  
AAACTTCTAAACTAATTGTTCTTTTGTGCTTTAATTTTTTAAAAATTACATTCTTCTGATGTA  
ACATGTGATACATACAAAAGAATATAGTTTAAATATGTATTGAAATAAAACACAATAAAATT

**FIGURE 52**

MPEMPEDMEQEEVNIPNRRVLVTGATGLLGRAVHKEFQQNNWHAVGCGFRRARPKFEQVNLLD  
SNAVHHIIHDFQPHVIVHCAAERRPDVVENQPDAAASQLNVDASGNLAKEAAAVGAFLIYISSD  
YVFDGTNPPYREEDIAPPLNLYGKTKLDGEKAVLENNLGAAVLRIPILYGEVEKLEESAVTVM  
FDKVQFSNKSANMDHWQQRFPTHVKDVATVCRQLAEKRMLDPSIKGTFHWSGNEQMTKYEMAC  
AIADAFNLPSSHLRPITDSPVLGAQRPRNAQLDCSKLETLGIGQRTPFIRIGIKESLWPFLLDK  
RWRQTVFH

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 105-127

**N-glycosylation site.**

amino acids 197-201

**N-myristoylation site.**

amino acids 303-309

**Short-chain dehydrogenases/reductases family proteins.**

amino acids 18-30





FIGURE 54

MLLIISLLLAAGLMHSDAGTSCPVLCTCRNQVVDCCSSQRLFSVPPDLPM DTRNLSLAHN RITAV  
PPGYLTCYME LQVLDLHNNSLMELPRGLFLHAKRLAHL DLSYNNF SHVPADMFQEAHGLVHID  
LSHNPWLRRVHPQAFQGLMQLRDL DLSYGGLAFLSLEALEGLPGLVTLQIGGNPWVCGCTMEP  
LLKWLNRNRIQRCTADSQLAECRGPP EVEGAPLFSLTEESFKACHLTTLTDDYLFIAFVG FVVS  
IASVATN FLLGITANCCHRWSKASEEEEEI

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 241-260

N-glycosylation sites.

amino acids 52-55, 81-84, 107-110

Tyrosine kinase phosphorylation site.

amino acids 148-154

N-myristoylation sites.

amino acids 11-15, 263-268

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 175-185

Leucine zipper pattern.

amino acids 77-98

## FIGURE 55

GGCTGCGCCACAGGCCGGCGGGGCCAGCAGCTGCGAACCCGCCCGCCACCACTGTTTCCGCGC  
CCGGGGACTTCCCCGGCGGGGCTCAGAAGTGTGGGTGCGTTCGCTTGGCTTCCCCCTGGCGTCA  
GCGACCCAGGGTAACCTCCTCCACTGCTGCGTGCCGTGCAGGCCTGCCTGTGTGAGAGCCACG  
TGTGCCGCGCTCTGGGCACAGCCTTGGAAGTCAAGGACCGCGACGGCAGCAGAGCAAACCT  
TACAGAAACATGAAGCCCTCAACCATCTGCTACTCAGTTATTTCGGGGCTGACGGCGGCTTCTA  
GAACATCCAGGTGTTCTGCAGATGCGAGAACTCATCCTGTAGTCACCAGATGGAGTCCCAAAC  
AGCCAAGCAGATGTAAGGCCTGTGCTGTGGCTCTGAGGCCCTGAATACAGAAGGGTCACTTTC  
TTAGTGGCCAAAGAGCAGTTGTTGACATTGATGTCTAATTATTGAACACGACCAGTCATTTTA  
CTGAGCTGCAGTGAGGAAACACTGACCATAGAAGATCAAGCCAAATGAGGGATTGCAAATTC  
CTGATTCTTTTGAATTAGGATTCCAGATGGGGGCCTCATTCTACAGCCCCCAACATTCCTAT  
AGCCGTTATCACTGCCATCACCCTGCCACCAGCATCTTCTTGAGATTCCACCCCTGCTCCC  
CAGAGACTTCCTGCTTTGAAAGTGAGCAGAAAGGAAGCTCTCAGAAAAATCTCTAGTGGTGGC  
TGCCGTGCGCTCCAGACAATCGGAATCCTGCCTTCACCACCATGGGGCTGGCTTTTTCTAAAGGT  
TTTGTTGGCGGGAGTGAGTTTCTCAGGATTCTTTATCCTCTTGTGGATTTTGCATCAGTGG  
GAAAACAAGAGGACAGAAGCCAAACTTTGTGATTATTTTGGCCGATGACATGGGGTGGGGTGA  
CCTGGGAGCAAACCTGGGCAGAAACAAGGACACTGCCAACCTTGATAAGATGGCTTCGGAGGG  
AATGAGGTTTGTGGATTTCCATGCAGCTGCCTCCACCTGCTCACCTCCCGGGCTTCCTTGCT  
CACCGGCCGGCTTGCCCTTCGCAATGGAGTCACACGCAACTTTGCAGTCACTTCTGTGGGAGG  
CCTTCCGCTCAACGAGACCACCTTGGCAGAGGTGCTGCAGCAGGCGGGTTACGTCACTGGGAT  
AATAGGCAAATGGCATCTTGGACACCACGGCTCTTATCACCCCAACTTCGGTGGTTTTGATTA  
CTACTTTTGAATCCCATATAGCCATGATATGGGCTGTACTGATACTCCAGGCTACAACCACCC  
TCCTTGTCAGCGTGTCCACAGGGTGATGGACCATCAAGGAACCTTCAAAGAGACTGTTACAC  
TGACGTGGCCCTCCCTCTTTATGAAAACCTCAACATTGTGGAGCAGCCGGTGAACCTTGAGCAG  
CCTTGCCCAGAAGTATGCTGAGAAAGCAACCCAGTTTCATCCAGCGTGCAAGCACCAGCGGGAG  
GCCCTTCCTGCTCTATGTGGCTCTGGCCACATGCACGTGCCCTTACCTGTGACTCAGCTACC  
AGCAGCGCCACGGGGCAGAAGCCTGTATGGTGCAGGGCTCTGGGAGATGGACAGTCTGGTGGG  
CCAGATCAAGGACAAAGTTGACCACACAGTGAAGGAAAACACATTCTCTGGTTTACAGGAGA  
CAATGGCCCGTGGGCTCAGAAGTGTGAGCTAGCGGGCAGTGTGGGTCCCTTCACTGGATTTTG  
GCAAACCTCGTCAAGGGGGAAAGTCCAGCCAAGCAGACGACCTGGGAAGGAGGGCACCGGGTCCC  
AGCACTGGCTTACTGGCCTGGCAGAGTTCCAGTTAATGTCACCAGCACTGCCTTGTTAAGCGT  
GCTGGACATTTTTTCCAACCTGTGGTAGCCCTGGCCCAGGCCAGCTTACCTCAAGGACGGCGCTT  
TGATGGTGTGGACGTCTCCGAGGTGCTCTTTGGCCGGTCAACAGCCTGGGCACAGGGTGCTGTT  
CCACCCCAACAGCGGGGCAGCTGGAGAGTTTGGAGCCCTGCAGACTGTCCGCTGGAGCGTTA  
CAAGGCCTTCTACATTACCGGTGGAGCCAGGGCGTGTGATGGGAGCATGGTGCCTGAGCTGCA  
GCATAAGTTTTCTCTGATTTTCAACCTGGAAGACGATACCGCAGAAGCTGTGCCCTAGAAAG  
AGGTGGTGCAGGAGTACCAGGCTGTGCTGCCCGAGGTGAGAAAGTTCTTGAGACAGTCCCTCA  
AGACATTGCCAACGACAACATCTCCAGCGCAGATTACACTCAGGACCCTTCAGTAACCTCCCTG  
CTGTAATCCCTACCAAATTGCCCTGCCGCTGTCAAGCCGCATAACAGACCAATTTTTTATTCCAC  
GAGGAGGAGTACCTGGAAATTAGGCAAGTTTGCTTCCAAATTTCATTTTTACCCTCTTTACAA  
ACACACGCTTTAGTTTAGTCTTGAGTTTGTGAGTTTGGAGTTAGCCTTGATATCCCTTCTGTA  
TCCTGTCCCCCTCCACGCCGACCCGAGAGCAGCTGAGCTGCGCTGGCTCTGGGCAGGGAGTG  
TGCCTTAATGGGAAGCACACGGGCTTTGGAGTCAGGCACAGGTGCCAGCTCCAGCTTTTGAAC  
TTGGGCAATTGTTTAAACCTAACCTGCAAGTTGATTTTGAGGGTTAAATAAAGGCATACATGAA  
AATGCCTGGCAACTTTAAAAA

**FIGURE 56**

MGWLFLKVLLAGVSFSGFLYPLVDFCISGKTRGQKPNFV I ILADDMGWGDLGANWAETKDTAN  
 LDKMASEGMRFVDFHAAASTCSPSRASLLTGRLGLRNGVTRNFAVTSVGGLPLNETTLAEVLQ  
 QAGYVTGIIGKWHLGHHGSYHPNFRGFDYYFGI PYSHDMGCTDTPGYNHPPCPACPQGDGPSR  
 NLQRDCYTDVALPLYENLNIVEQPVNLSLAQKYAEKATQFIQRASTSGRPFLLYVALAHMHV  
 PLPVTQLPAAPRGRSLYGAGLWEMDSL VGQIKDKVDHTVKENTFLWFTGDNGPWAQKCELAGS  
 VGPFTGFWQTRQGGSPAKQTTWEGGHRVPALAYWPGRVPVNVSTALLSVLDIFPTVVALAQA  
 SLPQGRFRFDGVDVSEVLFGRSQPGHRVLFHPNSGAAGEFGALQTVRLERYKAFYITGGARACD  
 GSMVPELQHKFPLIFNLEDDTAEAVPLERGGAEYQAVLPEVRKVLADVLQDIANDNISSADYT  
 QDPSVTPCCNPYQIACRCQAA

**Important features of the protein:****Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 353-373

**N-glycosylation sites.**

amino acids 117-120, 215-218, 356-359, 397-500

**N-myristoylation sites.**

amino acids 12-17, 33-38, 52-57, 97-102, 101-106, 113-118, 158-163, 328-333, 388-393, 418-423, 435-440, 436-441

**Amidation site.**

amino acids 382-385

**Sulfatases signature 2.**

amino acids 129-138

TGGACAAGACACCTCCAGGAGGCCAGCTCACAGCCACCGGTACCTTCTTCCAGGACAAGCTG  
GGGCCTCCATGGGCGCCTGAGGGCCAGGCGCCAGGGCCGTGGGCACGAGTATGGTGAGACACC  
AGCCCCCTGCAGTACTACGAGCCACAGCTGTGCCTCTCCTGCCTCACGGGCATCTACGGCTGCC  
GTTGGAAGCGCTACCAGCGCTCCCATGATGATACCACACCGGGCACAGCGCCATTCTCTGCATG  
TGGGGGCTGTGGCAGCAGTCACCATGCTCTCCTGGATCGTGGCAGGACAGTTCGCCCCGTGCAG  
AGCGGACCTCCTCCCAGGTGACCATTCTCTGTACCTTCTTCACCGTGGTGTGTTGCCCTCTACC  
TGGCCCCCTCTCACCATCTCCTCTCCCTGCATCATGGAGAAGAAAGACCTCGGCCCCAAGCCTC  
CTCTCATTGGCCACCGCGGGGCCCCCATGCTGGCTCCAGAGCACACGCTCATGTCTCTCCGGA  
AGGCCCTCGAGCAGAAGCTGTACGGGCTCCAGGCTGACATTACCATCAGCCTGGACGGCGTGC  
CCTTCCTCATGCATGACACCACCCTGCGGCGCACCAACGTGGAGGAGGAGTTCCCGGAGC  
TGGCCCGCAGGCCTGCCTCCATGCTTAACTGGACCACCCTGCAGAGACTCAACGCTGGCCAGT  
GGTTCCTGAAGACTGACCCCTTCTGGACAGCCAGCTCCCTGTCACCCTCCGACCACAGAGAGG  
CCCAGAACCAGTCCATCTGCAGCCTGGCAGAGCTCCTGGAGCTGGCCAAGGGCAATGCCACAC  
TGCTGCTCAACCTGCGTGACCCGCCCCGGGAGCACCCCTACCGCAGCAGTTTTATCAACGTGA  
CTCTGGAGGCCGTGCTGCACTCCGGCTTCCCCCAGCACCCAGGTCATGTGGCTGCCTAGCAGGC  
AGAGGCCCCCTGGTGCAGGAAGGTGGCTCCCGGCTTCCAACAGACATCAGGCTCCAAGGAGGCAG  
TCGCCAGCCTGCGGAGAGGCCACATCCAGCGGCTGAACCTGCGCTACACTCAGGTGTCCCGCC  
AGGAGCTCAGGGACTACGCGTCCCTGGAACCTGAGTGTGAACCTCTACACAGTCAACGCACCGT  
GGCTCTTCTCCCTGCTGTGGTGTGCGGGGGTCCCATCCGTACCTCTGACAACCTCCACACCC  
TGTCCCAGGTGCCTTCCCCCTCTGGATCATGCCCCCGGACGAGTACTGTCTCATGTGGGTCA  
CTGCCGACCTGGTCTCCTTACCCTCATCGTGGGCATCTTCGTGCTCCAGAAGTGGCGCCTGG  
GTGGCATAACGGAGCTACAACCCTGAGCAGATCATGCTGAGTGCTGCGGTGCGCCGGACACGCC  
GGGACGTGAGCATCATGAAGGAGAAGCTTATTTTCTCAGAGATCAGCGATGGTGTAGAGGTCT  
CCGATGTGCTCTCCGTATGTTTACAGACAACAGTTATGACACATATGCCAACAGCACCGCCACCC  
CTGTGGGCCCCCGAGGGGGTGGCAGCCACACCAAGACCCTCATAGAGCGGAGTGGGCGTTAGC  
TGAAGACATGTCTGTCCCACCTGTACCTGACACAGAAGCTGGGGAGCCTAGGAGAGCTGGTGG  
AAGTGTGTCTGAACTCGGAGTGCTCTGGGAGCGGGCTCCACAGCCTCCTTGTGGGCTCCAGCC  
CCTTGTCAGCCGCAGCCTCTCTTGAGGGGGACTCCCTGTCTCCTGAGGCCAGCTGGGCCAGG  
ACTCATCCTTTTTCAGATGCCCCCTGCAGGCCTGGGGCTCCTTCTGGGAAGTATGGGGCCTAGGG  
CTTGGTCCCCCTCTCTGAGGCCCTCTCCTGTATCCCCGACCTGGAAGCTTTGATGGGTGATGG  
GCCATGCCATACCCCTGTGGCAATGGAGTGCTGGATGCTCACCTGTGCCATCTGTCTCCTCCT  
GTCTGTGCCAGGAGGCACCTGAGTTCTCTGTCTGTTATCCTGCCCAAGGGCCTGGGCCGAGCC  
TCTACCTGAAGCAACTCTGCTCTTCTGTGAGTCTCAAAGCACAAAGAGGTTTCAGCCAGGAG  
GAAGCCAGCTGCAATGTGGAGACACGTCCTCCTCCCCAACCCACCTCATGCCACGCCAACCC  
CCTGCCCCAGGAGCGGGCCTGAGCCACGTCCCCCTAGGAGCAGCTGGAGATGGCCAAAAGAGTG  
AGCTCAGGACTACTGGATCCCATGCCCAGGTGTCCAGCAGACCTCAAGGCAGAAGGGTCACCT  
AACCCAGGAGTCCACAGACTGATGTGACCTCAGGTTCCACATCAGTGGCCACAGGGCAGGGC  
CCACCTGGTAGAAGTGTCTGGATATGGCCAGGGTGGGTGTGTGGCTAAGTGGGCCTGAACAG  
AGGGAACCTAGGGCCCTTGGCCAATGTGATTAAAGCTGCCATCTTGAAA

**FIGURE 58**

MVRHQPLQYYEPQLCLSLCTGIYGCRWKRYQRSHDDTTPGTAPFLHVGAVAAVTMLSWIVAGQ  
 FARAERTSSQVTILCTFFTVVFALYLAPLTISSPCIMEKKDLGPKPALIGHRGAPMLAPEHTL  
 MSFRKALEQKLYGLQADITISLDGVPFLMHD TTLRRTTNVEEEFPELARRPASMLNWTTLQRL  
 NAGQWFLKTDPFWTASSLSPSDHREAQNQSICSLAELLELAKGNATLLLNLRDPPREHPYRSS  
 FINVTLEAVLHSGFPQHQVMWLP SRQRPLVRKVAPGFQQTSGSKEAVASLRRGHIQRLNLRYT  
 QVSRQELRDYASWNLSVNLYTVNAPWLF SLLWCAGVPSVTS DNSHTLSQVPSPLWIMPPDEYC  
 LMWVTADLVSFTLIVGIFVLQKWRLGGIRSYNPEQIMLSAAVRRTSRDVSIMKEKLIFSEISD  
 GVEVSDVLSVCSDNSYDTYANSTATPVGPRGGGSHTKTLIERSGR

**Important features of the protein:****Signal peptide:**

amino acids 1-24

**Transmembrane domains:**

amino acids 47-61, 77-93, 335-350, 380-399

**N-glycosylation sites.**

amino acids 182-186, 217-221, 233-237, 255-259, 329-333, 462-466

**Tyrosine kinase phosphorylation site.**

amino acids 130-139

**N-myristoylation sites.**

amino acids 21-27, 48-54, 294-300, 404-410, 442-448, 473-479

**FIGURE 59**

CTTGAGCAAAACACAGCAGCCCGAGTGTTCCTCCAAAGGCCAAA**ATG**CTGAGAACGTCCTCACTCCTAA  
TCTGTGTGGTGGTCTGCATTGCCGGGGCCCCCTGGCTCTCTTCTGGCATTCTCTGCCTCTGCCT  
CATATTCTTGTTAGGCCAGGTGGGCTTGCTGCAGGGACACCCCCAGTGCCTGGATTACGGGCC  
CCCTTTCCAGCCCCCTCTGCACCTTGAGTTTTGCTCTGACTATGAGTCCTTCGGCTGCTGTGA  
TCAGCACAAGGACCGCCGCATCGCTGCCCGGTACTGGGACATCATGGAATATTTTGATCTGAA  
GAGACATGAGCTGTGTGGAGATTACATTAAAGACATCCTTTGCCAGGAGTGCTCGCCCTACGC  
AGCCACCTCTACGACGCCGAAAACACCCAGACGCCTCTCCGGAATCTCCCGGGCCTCTGCTC  
TGATTACTGCTCTGCCTTCCATTCTAACTGTCACTCAGCCATTTCCCTGCTGACCAATGACCC  
CGGCCTCCAGGAGTCTCATGGAAGGGACGGTACCCGCTTCTGCCACCTCCTGGACCTTCCTGA  
CAAGGACTATTGCTTCCCTAATGTCCTGAGGAACGACTATCTCAACCGCCACCTGGGCATGGT  
GGCCCAAGATCCTCAGGGCTGCCTGCAGCTCTGCCTGAGCGAGGTGGCCAACGGGCTGAGGAA  
CCCCGTCTCCATGGTCCATGCTGGGGACGGCACCCATCGCTTCTTTGTTGCCGAGCAGGTAGG  
AGTGGTGTGGGTCTACCTCCCTGATGGGAGTCGCTGGAGCAACCCTTCTGGACCTCAAGAA  
CATCGTGTGACCACCCCATGGATCGGGGATGAGAGAGGCTTCTTGGGGTTGGCTTTTCACCC  
CAAATTCGCCACAATCGCAAGTTCATATTTATTATTTCGTGCCTGGACAAGAAGAAGGTAGA  
AAAGATCCGAATTAGTGAGATGAAGGTTTTCTCGGGCTGATCCTAACAAAGCTGACCTGAAATC  
AGAGAGGGTCACTTGGAGATTGAAGAACCAGCCTCAAACCATAATGGCGGACAACCTTCTTTT  
TGGCCTGGATGGCTATATGTACATATTACTGGGGACGGGGGACAGGCTGGAGATCCCTTTGG  
CCTGTTTGGAAATGCTCAGAACAAAAGTTCCCTGCTGGGAAAAGTTTTAAGGATCGATGTGAA  
CAGGGCAGGCTCACATGGCAAGCGGTACCGAGTCCCCTCGGACAATCCATTTGTTTCTGAGCC  
AGGGGCCACCCCGCCATCTATGCCTATGGGATCAGGAACATGTGGCGTTGTGCTGTGGACCG  
AGGGGACCCCATCACGCGCCAGGGCCGAGGCCGGATATTCTGTGGGGACGTGGGCCAGAACAG  
GTTTGAAGAGGTTGACCTCATTTTGAAGGTTGAAACTATGGCTGGAGAGCAAAGGAAGGGTT  
TGCATGTTATGACAAAAAACTTTGTCACAATGCCTCTTTGGATGATGTTCTGCCAATCTATGC  
TTATGGCCATGCAGTGGGGAAGTCAGTCACTGGAGGTTATGTCTATCGTGGTTGTGAATCCCC  
AAATCTCAATGGCCTGTATATCTTTGGAGACTTCATGAGTGGTCGACTTATGGCTTTGCAGGA  
AGATAGAAAAACAAGAAATGGAAGAAGCAGGATCTTTGCCTGGGCAGCACACGTCCTGTGC  
CTTCCCAGGGCTGATCAGCACCCATAGCAAGTTCATCATCTCCTTTGCTGAAGATGAAGCAGG  
GGAGCTGTATTTCTGGCGACCTCTTACCCAAGTGCCTATGCACCACGTGGATCTATTTACAA  
GTTTGTGACCCCTCAAGGCGAGCACCCCCAGGCAAGTGCAAATACAAGCCAGTGGCCGTGAG  
AACCAAGAGTAAGCGGATCCCGTTCAGACCACTCGCCAAGACAGTCTTGGACTTGCTAAAGGA  
ACAATCAGAGAAAGCTGCTAGAAAATCTTCCAGTGCAACCTTAGCTTCTGGCCCAGCCCAGGG  
TTTGTCTGAGAAAGGCTCCTCCAAGAAGCTGGCTTCTCCTACAAGCAGCAAGAATACATTGCG  
AGGGCCTGGTACAAAGAAGAAAGCCAGAGTGGGGCCCCACGTCCGCCAGGGCAAGAGGAGGAA  
GAGCCTGAAAAGCCACAGTGGCAGGATGAGGCCATCAGCAGAGCAGAAGCGAGCTGGCAGAAG  
TCTCCCT**TGA**CCTATTGGTCAAGGTGGCCGACAGGGTGACGTGAGAGAGGAGAGCCACCTCAT  
CAAATGAAAGTCACTGCTGAATAAAGACCTTAGAAGTCTGGGAAGCCAGGGTAGAGGTGGGC  
AGGGCGGTTTTTCTCTCCCTGGGAAATCTTGCTGTCTACTGAATAAATAAATGCACCTTCTCT  
GTATGCAGTGCTTCTGTGGGAGACCATATCCCAGATTGCTGGTGCACCTGGGTTATGGTAAGC  
ACTAGTCCATGAGCCTGCTTGAATCACACTGGATGTCTCCGTTTTGTCTTGTAATGCCTAC  
AACCTGAGGTAATAAATCAACATTTGCTCA

**FIGURE 60**

MLRTSTPNLCGGLHCRAPWLSSGILCLCLIFLLGQVGLLQGHQPCLDYGPPFQPPLHLEFCSD  
 YESFGCCDQHKDRRIAARYWDIMEYFDLKRHELCDYIKDILCQECSPYAAHLYDAENTQTPL  
 RNLPGLCSDYCSAFHSNCHSAISLLTNDRGLQESHGRDGTFRFCHLLDLPDKDYCFPNVLRNDY  
 LNRHLGMVAQDPQGCLQLCLSEVANGLRNPVSMVHAGDGTHRFFVAEQVGVVWVYLPDGSRL  
 QPFLDLKNIVLTTPWIGDERGFLGLAFHPKFRHNRKFYIYYSCLDKKKVEKIRISEMKVSRAD  
 PNKADLKSERVILEIEEPASNHNGGQLLFGLDGYMYIFTGDGGQAGDPFGLFGNAQNKSSLLG  
 KVLRIDVNRAGSHGKRYRVPSDNPFVSEPGAHPAIYAYGIRNMWRCVDRGDPITRQGRGRIF  
 CGDVGQNRFEEDLILKGGNYGWRAKEGFACYDKKLCHNASLDDVLPYAYGHAVGKSVTGGY  
 VYRGCEPNLNGLYIFGDFMSGRLMALQEDRKNKKWKKQDLCLGSTTSCAFPGLISTHSHKFI  
 SFAEDEAGELYFLATSYPSAYAPRGSYKFVDPSSRRAPPGKCKYKVPVVRTKSKRIPFRPLAK  
 TVDLLKEQSEKAARKSSATLASGPAQGLSEKSSKKLASPTSSKNTLRGPGTKKKARVGPH  
 VRQGKRRKSLKSHSGMRPSAEQKRAGRSLP

**Important features of the protein:****Signal peptide:**

amino acids 1-41

**Transmembrane domain:**

amino acids 17-36

**N-glycosylation sites.**

amino acids 372-376, 480-484

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 645-649, 699-703

**Tyrosine kinase phosphorylation site.**

amino acids 81-89

**N-myristoylation sites.**

amino acids 11-17, 37-43, 156-162, 165-171, 357-363, 365-371,  
 368-374, 408-414, 459-465, 548-554, 557-563

**Amidation sites.**

amino acids 391-395, 696-700

**Cell attachment sequence.**

amino acids 428-431

**Leucine zipper pattern.**

amino acids 25-47



**FIGURE 61**

CTCCATTAAACCACCACCAGCTCCCCAAGCCACCCCTTCAGCC**ATG**AAGTTCCTGCTCCTGGT  
CTTGGCAGCCCTCGGATTCCCTGACCCAGGTGATCCCAGCCAGTGCAGGTGGGTCAAAATGTGT  
GAGTAACACCCCAGGATACTGCAGGACATGTTGCCACTGGGGGGAGACAGCATTGTTTCATGTG  
CAACGCTTCCAGAAAATGCTGCATCAGCTACTCCTTCCTGCCGAAGCCTGACCTACCACAGCT  
CATCGGTAACCACTGGCAATCAAGGAGAAGAAACACACAAAGGAAAGACAAGAAGCAACAAAC  
GACCGTAACATCA**TAA**TAACCACTGCTATCGCCTCCACCAACTCAGAGAAATATCATTTCAC  
AGTTCCAATTCCCTCCTACATTGCTGAGTACTAGCCAAGGCTCCTCTTTATGGGGCAGATATCT  
ATAGCCAACCCCAAACTTCTGTCTTCTATCATTCTGTCAATTCATCTAGTAACTAATTTGGAG  
TTTGTATCTATCTTACGAGAACAATCATCATGCAGATTCGTCCACAGGGGATCTGTCAGTTTG  
GGTCTCCAAATGAAAAATGTCAAGACAGAATTGGACATGCAAAGATTGACTGGGAGAACAC  
ACCTCTGATGGACAAAGGTGAGACAGAGCAGCCACAGGCAGGGAGAGCCTTCAGACTGCAACG  
CTGGCCTGATACGTGTCAAAGGAGAGAGGGATAGAGGAGGATTGAATAGAAGGAGACTAAGAC  
TGCAGCTCTAAGAAAGTCTCAGCCAAACAGATGGGGAGGCCCAAAGCAAGGCTTGCCCCCTCAG  
AGGAGCTCACGCAGGGCAGGAATAGCCAGGTTCTCATATCCCAGGGGTTTCAGACTTGGCTGAG  
AACAGCCCCCTGGAGAACATGGGGTGACTGCTACCATAGGTCTGGAAGTATGAGGCTGTCCACC  
AACTATCCCCTTGAAGCAAGTTCTCTTGAAAGGAAATCTAAACAGTGCACCCCCATGGCTGCC  
ACGGAGTATAAGGAGGGAGAGAAAGGAGCTGAAAGTCTAGGTTTGGCCAGCTAGGTAGACTGA  
CTTGTGAGGTATTTATTTATTCATTTGAGTAACAAAGCAGACAGAATACATAGCCACCATTGG  
TAGTACACCCCCAAAAGCAAGGATGGCATGATGCTGGTGACTCAAACGTGCCTACTCATGGTGT  
CAAATTGGCATAATCCTCTTGGAAGCTGTGTGGAATAAGCACAGAGAAGCAGAACTCTAAT  
TGCTTAATCCACTAAACATTACTTCTGGGAATTGGCTCATCATAAATTATCCAAGAGAAAGCA  
CAAAGTTATGGGCACAAAGGTTTTCCATATAATATTATTTAAAATGCTGAGAAAATGAAAAAA  
TCTAAATGGTGAAATATATACTAATGCCATCTATAAATACAAACAAATAGAATGTTTATAGAA  
TAATGGAACATAATAACATTATTCAAATTTGCATTTATGCTATAGTTGTCAAATTTGTCTCCT  
TATATGATACAAAACCTCATGAAAATTATGACTTTTTTGTGGTTGGAAAGCAGAATTATGCA  
TAAATTTCCCTCTTACAGTTCGATGCCCATAGTTTTATATAACATTTATTTGACACGTACTGA  
CTTCTATCTGAGAAGAACAAACCAAAACACTCAGGCCTAAATAATTA AAAACGGTCCTAAAAA  
CTAGCAAACCAGATAAGAAAAGATGTTAATGCCCATTCCTAACTTATGTCTTAGACCAAAAT  
TAATTCTAGATGGTTTTAAAATGACAGTGTAAGTAAAGTATTAAAAGATTGTGTGGTCAAA  
TATTCAATTTAAGAGCAAGGAAATTCTTATAAATATAACAATAGAGGCAGAACTCATGTAAGA  
ATAAATTGATTAGGTGGTATTAAATATTAAGTTCTTATGTATGTCAAAGATATCATTTTTGAA  
ATTCATCCATCTTATTGGGTATTGCAGGAGTTCATTCCTTTTTGTTTATAAATACTCTTCCGT  
CATATGAATAGTATTCATTTGTATACTGGTTTTGTTGATGGACATTTGGGTTGTTCCCAGTTTA  
TGGCTATTACAAATAAAGCTTCTATGAACATTTATGTACA





**FIGURE 64**

MEATGTWALLLALALLLLTLALSGTRARGHLPPGPPTPLPLLGNLLQLRPGALYSGLMRLSKK  
 YGPVFTIYLGWRPVVVLVGQEAVERALGGQAEFEFSGRGTVMLEGTFDGHGVFFSNGERWRQ  
 LRKFTMLALRDLGMGKREGEELIQAEARCLVETFQGTGRPFDPSTLLAQATSNVVCSTLFLGL  
 RFSYEDKEFQAVVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHHVSTLAAFTVRQV  
 QQHQNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLLFAGTMTVSTTVGYTL  
 LLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPTYDAVLHEAQRLALVPMGIPRTLMRT  
 TRFRGYTLPQGTEVFPLLSILHDPNIFKHPEEFNPDRFLDADGRFRKHEAFLPFSLGKRVCL  
 GEGLAKAELFLFFTTILQAFSLSPCPPDTLSLKPTVSGLFNIIPPAFQLQVRPTDLHSTTQTR

**Important features of the protein:****Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 294-313

**Glycosaminoglycan attachment site.**

amino acids 99-103

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 128-132

**N-myristoylation sites.**

amino acids 51-57, 109-115, 115-121, 188-194, 207-213, 257-263,  
 284-290, 339-345, 370-376, 444-450

**Amidation sites.**

amino acids 140-144, 435-439

**Leucine zipper pattern.**

amino acids 32-54, 39-61

**Cytochrome P450 cysteine heme-iron ligand signature.**

amino acids 433-443

## FIGURE 65

CGGACGCGTGGGGCCGTATGCGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCC  
CCAGCCCCCTGCTCCTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGA  
AGACCCGCCAGGTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTC  
ATATACGGGCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGG  
CAGTGGTAATGGTGGCCACCAACACCCCCCACAGCACCTGAGCATCAACTGGAGCCTCCTGC  
TATCCCCTGAGCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTCTTCTG  
CCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCAAAGC  
CTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATCACTGATT  
CATTTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGACCCTACCAGGA  
CTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCCAGCCGACCAGCCC  
AACCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTGGCCCTGATTGGAGCCT  
CTCCCCGGGGAAACCGTTCCCTGTTTGGGCTGGAGGTAGCCACATTGGGCCAGGGCCCTGACT  
GCCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATATGCACCGGCCGTCTTCCAGTTGG  
ACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCACAGTGGCGACCAGTGGCTTACTCCC  
AGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGCCAAGCTTCCCCTCTTCATCCTGCCTTAG  
CATACTCTCTTCCCCAGTCACCCATTGTCCGAGCCTTCTTTGGGTCCCAGAATAACTTCTGTG  
CCTTCAATCTGACGTTTCGGGGCTTCCACAGGCCCTGGCTATTGGGACCAACACTACCTCAGCT  
GGTCGATGCTCCTGGGTGTGGGCTTCCCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCA  
TCATGGCAGTGGCCCTGGGTGCCCCAGGGCTCATGCTGCTAGGGGGCGGCTTG GTTCTGCTGC  
TGCACCACAAGAAGTACTCAGAGTACCAGTCCATAAATTAAGGCCCGCTCTCTGGAGGGAAGG  
ACATTACTGAACCTGTCTTGCTGTGCCTCGAAACTCTGGAGGTTGGAGCATCAAGTTCAGCC  
GGCCCCCTTCACTCCCCCATCTTGCTTTTCTGTGGAACCTCAGAGGCCAGCCTCGACTTCCTGG  
AGACCCCCAGGTGGGGCTTCCCTTCATACTTTGTTGGGGGACTTTGGAGGCGGGCAGGGGACAG  
GGCTATTGATAAGGTCCCCTTGGTGTTGCCTTCTTGCATCTCCACACATTTCCCTTGGATGGG  
ACTTGCAGGCCTAAATGAGAGGCATTCTGACTGGTTGGCTGCCCTGGAAGGCAAGAAAATAGA  
TTTATTTTTTTTTTACAGGGAAAAAAAAAAAAA



**FIGURE 67**

CGGGACAGGCGCGTGAGGCCACAACACATGCGTGTATCTTGCTTGGGCTATCTTCCCTGCTCTGCCACGCCGGGT  
CTGGAGAAGGGGTTTCAGCCCCAGGACATTTACTGAGAGTGGCGGAATATTGGGAGCCGCGCATGTTCCCCCTTCG  
GGCCCTGTGGTTGGTCTGGGCGCTTCTAGGAGTGGCCGGATCATGCCCGAGCCGTGCGCCTGCGTGGACAAGTA  
CGCTCACCAGTTCGCGGACTGCGCTTACAAAGAGTTGCGTGAGGTGCCGGAAGGACTGCCTGCCAACGTGACGAC  
GCTTAGTCTGTCCGGAACAAGATCACTGTGCTGCGGCGCGGGGCTTCGCCGACGTACACAGGTCACGTGCT  
GTGGCTGGCGCACAATGAGGTGCGCACCGTGAGCCAGGCGCACTGGCCGTGCTGAGTCAGCTCAAGAACCTCGA  
TCTGAGCCACAACCTTCATATCCAGCTTTCCGTGGAGCGACCTGCGCAACCTGAGCGCGCTGCAGCTGCTCAAAT  
GAACCACAACCGCCTGGGCTCTCTGCCCCGGGACGCACTCGGTGCGCTACCCGACCTGCGTTCCCTGCGCATCAA  
CAACAACCGGCTGCGTACGCTGGCGCCTGGCACCTTCGACGCGCTTAGCGCGCTGTCACTTGAACCTCTATCA  
CAATCCCTTCCACTGCGGCTGCGGCCTTGTGTGGCTGCAGGCCGAGCCCGGGTGTCCCTTACCCGA  
GCCGACTCCATTGCTTGTGCTCGCCTCCCGCGCTGCAGGGGGTGGCGGTGTACCGCCTGCCCCCCTGCCCTG  
TGCACCGCCAGCGTGCATCTGAGTGCCGAGCCACCGCTTGAAGCACCCGCGACCCCACTGCGCGCAGGACTGGC  
GTTCTGTGTACACTGCATCGCCGACGGCCACCTACGCTCGCCTGCAATGGCAACTTCAGATCCCCGGTGGCAC  
CGTAGTCTTAGAGCCACCGGTTCTGAGCGGGGAGGACGACGGGGTGGGGCGGAGGAAGGAGAGGGGAGAAGGAGA  
TGGGATTTGTGATGACGCAACCAAGCCAAACGCGCACTCCAGCACCCGCTTGGCCGGCGCCCCAGCCACACC  
GCGCTTCTGGCCCTCGCAAATGGCTCCCTGTTGGTCCCCCTCTGAGTGCCAAGGAGGCGGGGCTACACTTG  
CCGTGCACACAATGAGCTGGGCGCAACTCTACGTCAATACGCGTGGCGGTGGCAGCAACCGGGCCCCCAACA  
CGCGCCTGGCGCGGGGGAGAACCCGACGGACAGGCCCCGACCTCTGAGCGCAAGTCCACAGCCAAGGGCCGGG  
CAACAGCGTCTGCTTCCAAACCCGAGGGCAAAATCAAAGGCCAAGGCCTGGCCAAGGTGAGATTCTCGGGGA  
GACCGAGACGGAGCGGAGGAGGACACAAGTGAGGGAGAGGAGGCCGAAGACCAGATCCTCGCGGACCCGGCGGA  
GGAGCAGCGCTGTGGCAACGGGACCCCTCTCGGTACGTTTCTAACCACGCGTTCAACCAGAGCGCAGAGCTCAA  
GCCGACGTCTTCGAGCTGGGCGTCACTCGCGCTGGATGTGGCGGAGCGCGAGGCGCGGGTGCAGCTGACTCGCT  
GGCTGCGCGCTGGGGCCCTGGGCCCCGGCGGGCTGGCGGAGCCCCGCGACCCGGGCGGCGACCCCTGCGCCTACT  
CTATCTGTGTCCAGCGGGGGGGCGGCGGCGAGTGCAGTGGTCCCGCTAGAGGAAGGCGTCAACGCTACTGGTT  
CCGCGGCCTGCGGCCGGGTACCAACTACTCCGTGTGCTGGCGCTGGCGGGCGAAGCCTGCCACGTGCAAGTGGT  
GTTTTCCACCAAGAAGGAGCTCCCATCGCTGCTGGTCACTAGTGGCAGTGAGCGTATTCCTCCTGGTGTGGCCAC  
AGTGCCCTTCTGGGCGCGCCTGCTGCCATCTGCTGGCTAAACACCCGGGCAAGCCCTACCGTCTGATCCTGCG  
GCCTCAGGCCCCCTGACCCTATGGAGAAGCGCATCGCCGCGACTTCGACCCGCGTGCTTCGTACCTCGAGTCCGA  
GAAAGCTACCCGGCAGGCGGCGAGGCGGGCGGCGAGGAGCAGAGGACGTGCAGGGGGAGGGCCTTGATGAAGA  
CGCGGAGCAGGGAGACCCAAGTGGGGACCTGCAGAGAGAGGAGAGCCTGGCGGCCTGCTCACTGGTGGAGTCCCA  
GTCCAAGGCCAACCAAGAGGAGTTCGAGGCGGGCTCTGAGTACAGCGATCGGCTGCCCTGGGCGCGGAGGCGGT  
CAACATCGCCAGGAGATTAATGGCAACTACAGGACAGGCGAGGCTGAACCTCCGCCCGTCCGGCCCCGCCATT  
CCCGACCTCCACCTAGGGTGCCTGGGAGCAGCAGTCTAGGGCTGGCAGGACTTATGTCCCCCGTCCCCAACCTTC  
ACCTACTCCTCCCCCTTACTACTCCCCAACCTTGACTACCAGGCACTTCTATTAGGGAGTGGGCCATTTCACCA  
GTCCCTGCTACCCACGGCTGCCATTCTCCCTGCGGGCTGAATCCCCTTCCCCGCCAAGCACAGTGTCTTACTTAC  
CCCATGCAAGACTCCACCCGACAGCGGTGGGCGATATCTATGTCCCTCCATTCCCGTCCGCGATTATCTGCGAAAT  
CCACCCCGCAGCCCGCCCCACCGTGGGCTCTGGAGCCAGAGGAAACGAGCGAAGACTTTGGAAACCTCGCGGTAA  
CGCGGTGGTTTTCGGGGGCCAGCCAAGGCCAGTGGAGTGTGTGGGGTCCACCTCGACCCCTCCTCCTCCCTTTC  
TTTCTTTCTTTTTTTTTTTTATTTTTTAAATTTATTTATTTATTTATTTATTTTATTTTACGGAGTCTTGGTCTGTGCG  
CAGGCTGGAGTGCAGTGGCGGATCTCGGCTCACTGCATCTTCCGCTCCCGGGTTCAAGCGATTCTCCTGCCTC  
AGCCTGCCTAGTAGCTGGGACTACAGGCGCGCGCCACCACGACCAGCTAATTTCTTCTATTTTTAGTAGAGACGG  
GGTTTCACCATGTTGGCCAGGATGGTCTGGATCTCTTGACCTCAGGTGATCCATCTGCCTCGGCCTCTCAAAGTG  
CTGGGATTACAGGCGTGAGGCACCGCGCCCGGCCCTCCTCCCTTTCAATCCCTACTCCAGAAGCCGGGATTCTG  
TGGCAACCCCTAGTTTTTATTTCCAAAGCCTCCTGCCGGCAGGGAACCAATCCTTCTGTCTCTCCACCCCCACC  
CCACTTCTGGCCAGTTGGAGTCCAGCCCGGTGCCTGGGGCGCCTTTCAGCTCCGCGCTCAGATTTTCTGTTTTT  
GTTGTTTTTCAAAGACAGCGACATTTCCGGTCTGGTGTAAACACCCCTTCCCAGCCTCTGGGAAAATCGAGTGTG  
TGTGTCCGGGGGTAGGGAGGGAATGCGTTTTCTGTGCTCTCTCTCCTAACTTAAAGCGCCGAGGACCGCGCGCC  
CCTTGGCGGCTGAGCCTGTGGACTTGGTGCAGGGCAATTTCTGTTGTCCGTGTGTTGGGCTTTCGGGAGGTCTGT  
GCGCCCAACAGCGCGCTCCCGCGGCTCCACCCGACCCAGACCTAGCTGGAAAGCGCCGGAGGCGGAGGAAGCT  
GACTGTGGCCTCCCGGGCGCGGCTCTTGAGGGCTCGCGCCCTAGTTTCGCACAAAGCCTGCTCGTGACTGTGC  
GACTGTGCGACGGGATCCGGATGGAGCCGAGCCCCCTCGCTCCTCGCTCTCGCTCCTCGCTCGCTCGCTCGCTCGCT  
CCGCCCCCTGCTTCCGGCGGGAATCGTGTGTGCCCCGGCGTGTAGTCCCTGACAAGCGTGCCCTGTAGGAGAAAAGTC  
TGTGTCTGTGAAGTGTGACCGTGTAGTGTAGGGGGGCGGGCGGGGGGGCGGATGGGCGGGAGGGAGGGGAAGGG  
GAGGGGCGGGCGCGCGGCTCGGGGCGGGGTTCTTTTTTCCATTTTGAAGAAAGCGTCCGGGTTGGGGTGGGG  
GGAGTTTCAGTCTCGGGATCAGCCCTCTCCGCGAAGCGCAGCACAAAGCGGGGCTGGGACGGAGTAGCCCCC  
GGAGCCCGTGCCCTTTCTAAACGCGTCTGTATGCACTCAATAAAACAATCGATTTGAAA

**FIGURE 68**

MFPLRALWLWVWALLGVAGSCPEPCACVDKYAHQFADCAVKELREVPEGLPANVTTLSSLANKI  
 TVLRRGAFADVTQVTSLWLAHNEVRTVEPGALAVLSQLKNLDLSHNFISSFPWSDLRNLSALQ  
 LLKMNHNRLGSLPRDALGALPDLRSLRINNRLRTLAPGTFDALSALSHLQLYHNPFHCGCGL  
 VWLQAWAASTRVSLPEPDSIACASPPALQGVVYRLPALPCAPPSVHLSAEPPEAPGTPLRA  
 GLAFVLHCIADGHPTPRLQWQLQIPGGTVVLEPPVLSGEDDGVGAEEGEGEGDGDLLTQTQAO  
 TPTPAPAWPAPPATPRFLALANGSLLVPLLSAKEAGVYTCRAHNELGANSTSIKVAVAATGPP  
 KHAPGAGGEPDQGAPTSEKSTAKGRGNSVLPSKPEGKIKGQGLAKVSILGETETEPEDTSE  
 GEEAEDQILADPAEEQRCGNGDPSRYVSNHAFNQSAELKPHVFELGVIALDVAEREARVQLTP  
 LAARWGPGPGGAGGAPRPGRRPLRLLYLCPAGGGAQVQSRVEEGVNAYWFRGLRPGTNY SVC  
 LALAGEACHVQVVFSTKKELPSLLVIVAVSVFLLVLATVPLLGAACCHLLAKHPGKPYRLILR  
 PQAPDPMEKRIAADFDPRASYLESEKSY PAGGEAGGEEPEDVQGEGLDEDAEQGDPSGDLQRE  
 ESLAACSLVESQSKANQEEFEAGSEYSDRLPLGAEAVNIAQEINGNYRQTAG

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 587-610

**N-glycosylation sites.**

amino acids 52-55, 121-124, 337-340, 364-367, 474-477, 563-566

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 397-400

**Casein kinase II phosphorylation sites.**

amino acids 19-23, 202-205, 289-292, 246-249, 411-414, 431-434,  
 433-436, 440-443, 544-547, 583-586, 650-653, 700-703

**N-myristoylation sites.**

amino acids 15-20, 48-53, 165-170, 296-301, 351-356, 362-367,  
 390-395, 419-424, 514-519, 536-541, 557-562, 561-566, 610-615,  
 661-666, 716-721

**Amidation site.**

amino acids 522-525

**Prokaryotic membrane lipoprotein lipid attachment sites.**

amino acids 10-20, 603-613



GGCGGCGGGAGCAGCAGCAAGAGGGGGCGGCAGGAGGATCTCTCAGGCTGCCGGTGGGAAGGCGTGGG  
 CGACCCGGTGTGTGGCGCGCCAGAGCCCCGCGTTTCAGCCCTAGGGAAGGAAGCAGTTGAG  
 GGAAGTTCTCCATGAATGTACGTCACAATGATGATGACCGACCAAATCCCTCTGGAAGTGGCA  
 CCATTGCTGAACGGAGAGGTAGCCATGATGCCCCACTTGGTGAATGGAGATGCAGCTCAGCAT  
 GTTATTCTCGTTCAAGTTAATCCAGGTGAGACTTTCACAATAAGAGCAGAGGATGGAACACTT  
 CAGTGCATTCAAGGACCTGCTGAAGTTCCCATGATGTCACCCAATGGATCCATTTCCTCCCAT  
 CATGTGCCTCCAGGTTATATCTCACAGGTGATTGAAGATAGTACTGGAGTCCGCCGGGTGGTG  
 GTCACACCCCAGTCTCCTGAGTGTTATCCCCCAAGCTACCCCTCAGCCATGTCTCCAACCCAT  
 CATCTCCCTCCCTATCTGACTCACCATCCACATTTTATTTCATAACTCACACACGGCTTACTAC  
 CCACCTGTTACCGGACCTGGAGATATGCCGCCTCAGTTTTTTTCCCCAGCATCATCTTCCCCAC  
 ACAATATATGGTGAGCAAGAAATTATACCATTTTTATGGAATGTCAAGCTACATCACCCGAGAA  
 GACCAGTACAGCAAGCCTCCGCACAAAAAACTGAAAGACCGCCAGATCGATCGCCAGAACCGC  
 CTCAACAGCCCTCCTTCTTCTATCTACAAAAGCAGCTGCACAACAGTATACAATGGCTATGGG  
 AAGGGCCATAGTGGTGAAGTGGCGGAGGCGGCAGCGGTAGTGGTCCCGGAATTAAGAAAACA  
 GAGCGACGAGCAAGAAGCAGCCCAAAGTGAATGATTGAGACTTGCAAGAATATGAGTTGGAA  
 GTAAAGAGGGTGCAAGACATTCTTTTCGGGAATAGAGAAAACACAGGTTTCTAATATTCAGGCA  
 AGAGCAGTTGTGTTGTCTGGGCTCCCCCTGTTGGACTTTCTGTGGACCCACAGTGGTCTT  
 TCCTTCCCCTACAGTTACGAGGTGGCCTTATCAGACAAAGGACGAGATGGAAAATACAAGATA  
 ATTTACAGTGGAGAAGAATTAGAATGTAACTGAAAGATCTTAGACCAGCAACAGATTATCAT  
 GTGAGGGTGTATGCCATGTACAATTCCGTAAAGGGATCCTGCTCCGAGCCTGTTAGCTTCACC  
 ACCACAGCTGTGCACCCGAGTGTCTTTCCCCCTAAGCTGGCACATAGGAGCAAAAGTTCA  
 CTAACCCTGCAGTGGAAGGCACCAATTGACAACGGTTCAAAAATCACCAACTACCTTTTAGAG  
 TGGGATGAGGGAAAAAGAAATAGTGGTTTCAGACAGTGCTTCTTCGGGAGCCAGAAGCACTGC  
 AAGTTGACAAAGCTTTGTCCGGCAATGGGGTACACATTCAGGCTGGCCGCTCGAAACGACATT  
 GGCACCAGTGGTTATAGCCAAGAGGTGGTGTGCTACACATTAGGAAATATCCCTCAGATGCCT  
 TCTGCACCAAGGCTGGTTCGAGCTGGCATCACATGGGTACGTTGCAGTGGAGTAAGCCAGAA  
 GGCTGTTACCCGAGGAAGTGATCACCTACACCTTGAAATTCAGGAGGATGAAAATGATAAC  
 CTTTTCCACCCAAAAATACACTGGAGAGGATTTAACCTGTACTGTGAAAAATCTCAAAGAAGC  
 ACACAGTATAAATTCAGGCTGACTGCTTCTAATACGGAAGGAAAAAGCTGTCCAAGCGAAGTT  
 CTTGTTTGTACGACGAGTCTTGACAGGCCTGGACCTCCTACCAGACCGCTTGTCAAAGGCCCA  
 GTTACATCTCATGGCTTTAGTGTCAAATGGGATCCCCCTAAGGACAATGGTGGTTTCAGAAATC  
 CTCAAGTACTTGCTAGAGATTACTGATGGAAATTCTGAAGGTGAAGTTTTTGGCAATTGTTTT  
 ATTCAAATCCAATTAGCAAGCTCTGTTTTCTAATATAGTAAATGTCTTTATAGTAATAGTGAGT  
 AATCATTAATTCTAAAGATAGAATTATTATTACAATAAACAACTTTAGTACATATTGGCAG  
 TTTTTCTATTTCAAACACAGCACCAGAGATCAGAGTCTACTTGAACTTACATTTGTGTTATT  
 TAACAATTTTTCTGTATCTTTTTTCATTGGTGTTTTGTTTTGTTTTATCTTTTGTTTTGT  
 TTGGTTTGGTTTGTTTTGTTTTGTTTTGTGAGATACGATCTCTGTCACACAGGCTGGAGGGC  
 AGTGGCACAGACATGGCCCATTCAGTCTCAGACTCCTGGGCTTAAGTGACTCTTCTGCCACA  
 GAAGATGAGGAAGAATACATTTTTTCATAGTGATGGGGTCTCACTATGTTATCTAGGCTGGTCT  
 CAACTCCTGGCCTCAAGCAACCCTCCACCTTGGCCTCCCAAAGTGCTGGGACTATAGACATG  
 AATCACACACTCAGCTTCCATGTCTTTTTATGAACTAGGGTTCCATAATTAATCAGATAAATT  
 TGGTATTTTCATCTCCTAACTTGCCATATGTTTTCTGGAAATCTTATAAGCAGCCGAGAGTG  
 GTGCTCACGCTGTAGTCCCAGCACTTTGGGAGGCTGAGGTGGGTGGTCAGGAGATCAAGACC  
 ATGCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAAATACAAAATTAGCTGGGTGTGGTG  
 GCAGGCACCTGTAGTCCCAGCTACTTTGGGAGGCTGAGGCAGAAGAATTGCTTGAACCCAGCAG  
 GCGGAGGTTGCAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGTGACAGAGTGAGACTC  
 TGTCTCAAAAAAAAAAAAAA

**FIGURE 70**

MMMTDQIPLELPPLLNGEVAMMPHLVNGDAAQHVLVQVNPGETFTIRAEDGTLQCIQGPAEV  
 PMMSPNGSIPPIHVPPGYISQVIEDSTGVRRVVVTPQSPECYPPSYPSAMSPTHHLPPYLTHH  
 PHFIHNSHTAYYPPVTGPGDMPPQFFPQHHLPHHTIYGEQEIIIPFYGMSSYITREDQYSKPPHK  
 KLKDRQIDRQNRNLNSPPSSIIYKSSCTTVYNGYGKGHSGGSGGGSGSGPGIKKTERRARSSPK  
 SNDSDLQEYELEVKRVQDILSGIEKPQVSNIQARAVVLSWAPPVGLSCGPHSGLSFPYSYEVA  
 LSDKGRDGKYKIIYSGEELEC�LKDLRPATDYHVRVYAMNSVKGSCSEPVSFTTHSCAPECP  
 FPPKLAHRSSKSSLTLQWKAPIDNGSKITNYLLEWDEGKRNSGFRQCFFGSQKHCKLTKLCPAM  
 GYTFRLAARNDIGTSGYSQEVVCYTLGNIPQMPSAPRLVRAGITWVTLQWSKPEGCSPEEVIT  
 YTLEIQEDENDNLFHPKYTGEDLTCTVKNLKRSTQYKFRLTASNTEGKSCPSEVLVCTTSPDR  
 GPPPTRPLVKGPVTSHGFSVKWDPPKDNNGSEILKYLLEITDGNSEGEVFGNCFIQIQ

**Important features of the protein:****N-glycosylation sites.**

amino acids 69-73, 254-258, 401-405

**Glycosaminoglycan attachment sites.**

amino acids 229-233, 234-238, 236-240

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 416-420, 535-539

**Tyrosine kinase phosphorylation site.**

amino acids 319-326

**N-myristoylation sites.**

amino acids 52-58, 227-233, 228-234, 230-236, 231-237, 232-238,  
 235-241, 239-245, 402-408, 610-616

**Amidation site.**

amino acids 414-418

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 290-301

**ATP/GTP-binding site motif A (P-loop).**

amino acids 546-554

**CUB domain proteins profile.**

amino acids 294-301

**FIGURE 71**

AAGTCATTCAAGTGGATGTGATCTTGGCTCACAGGGGACG**ATGT**CAAGCTCTTCCTGGCTCCTTCTCAGCCTTGTT  
 GCTGTAACGTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATTTTTGGACAAGTTTAACCACGAAGCCGAA  
 GACCTGTTCTATCAAAGTTCAGTTGCTTCTTGGAAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATG  
 AATAATGCTGGGGACAAATGGTCTGCCTTTTTAAAGGAACAGTCCACACTTGCCCAAATGTATCCACTACAAGAA  
 ATTCAGAATCTCACAGTCAAGCTTCAGCTGCAGGCTCTTCAGCAAAATGGGTCTTCAGTGCTCTCAGAAGACAAG  
 AGCAAAACGGTTGAACACAATCTAAATACAATGAGCACCATTACAGTACTGGAAAAGTTTGTAACCCAGATAAT  
 CCACAAGAATGCTTATTACTTGAACCAGGTTTGAATGAAATAATGGCAAAACAGTTTAGACTACAATGAGAGGCTC  
 TGGGCTTGGGAAAGCTGGAGATCTGAGGTGCGCAAGCAGCTGAGGCCATTATATGAAGAGTATGTGGTCTTGA  
 AATGAGATGGCAAGAGCAAATCATTATGAGGACTATGGGGATTATTGGAGAGGAGACTATGAAGTAAATGGGGTA  
 GATGGCTATGACTACAGCCGCGGCCAGTTGATTGAAGATGTGGAACATACCTTTGAAGAGATTAAACCATTATAT  
 GAACATCTTCATGCCTATGTGAGGGCAAAGTTGATGAATGCCTATCCTTCTATATCAGTCCAATTGGATGCCTC  
 CCTGCTCATTTGCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACTCTTTGACAGTTCCCTTTGGACAG  
 AAACCAAACATAGATGTTACTGATGCAATGGTGGACCAGGCCTGGGATGCACAGAGAATATTCAAGGAGGCCGAG  
 AAGTTCTTTGTATCTGTTGGTCTTCCTAATATGACTCAAGGATTCTGGGAAAATTCCATGCTAACGGACCCAGGA  
 AATGTTTCAGAAAGCAGTCTGCCATCCCACAGCTTGGGACCTGGGGAAGGGCGACTTCAGGATCCTTATGTGCACA  
 AAGGTGACAATGGACGACTTCTGACAGCTCATCATGAGATGGGGCATATCCAGTATGATATGGCATATGCTGCA  
 CAACCTTTTCTGCTAAGAAATGGAGCTAATGAAGGATTCCATGAAGCTGTTGGGGAATCATGTCACTTTCTGCA  
 GCCACACCTAAGCATTTAAATCCATTGGTCTTCTGTCAACCGATTTTCAAGAAGACAATGAAACAGAAATAAAC  
 TTCCTGCTCAAACAAGCACTCACGATTGTTGGGACTCTGCCATTTACTTACATGTTAGAGAAGTGGAGGTGGATG  
 GTCCTTAAAGGGGAAATTTCCCAAAGACCAGTGGATGAAAAGTGGTGGGAGATGAAGCGAGAGATAGTTGGGGTG  
 GTGGAAACCTGTGCCCCATGATGAAACATACTGTGACCCCGCATCTCTGTTCCATGTTTCTGATGATTACTCATT  
 ATTCGATATTACACAAGGACCCTTTACCAATTCCAGTTTCAAGAAGCACTTTGTCAAGCAGCTAAACATGAAGGC  
 CCTCTGCACAAATGTGACATCTCAAACCTCTACAGAAGCTGGACAGAACTGTTG**TAA**GAAATACCTCAAATGTT  
 GAACCTCTCCTAGTATTTCAGTATTACTCATTTCCATGCCTAGGTTTGTATTTGATTTCTTTGTTCTAAAAAGAAA  
 ATTTTATGGCCTCAAATGTCCTCATTTACAAACCAAACATTTAATTTGTGGTCAGACAGGAACCTAGACCATAC  
 AACAAATTGGGTGGGCCACCTCTTTTCTCCCTATCATACTACAGCCCTCTCTTCTGGTAATTGGAAGGAAAGAG  
 CGGTTTAGGGTGGAAATATATCTGTTAATATGCATTCTTTCTTATCTGCCAGAAGCAAATTTAGCCAAGTCAAAG  
 AGAAGAAACCATAGATCATAGATGTAATATATGTACATCTGGAACCCCTCAAAGGCCCTGAACCCCTTTTTT  
 TGTGTAGCAATATGCTGAGGCTTGGAAATCAGAACCCTGGACCCTAGCATTGGAAAATGTTGTAGGAGCAAGAA  
 CATGAATGTAAGGCCACTGCTCAACTACTTTGAGCCCTTATTTACCTGGCTGAAAGACCAGAACAAGAATTCCTT  
 TGTGGGATGGAGTACCGACTGGAGTCCATATGCAGACCCAAAGCATCAAAGTGAGGATAAGCCTAAATCAGCTC  
 TTGGAGATAAAGCATATGAATGGAACGACATGAATGTACCTGTTCCGATCATCTGTTGCATATGCTATGAGGC  
 AGTACTTTTTTAAAGTAAAAAATCAGATGATTCTTTTTGGGGAGGAGGATGTGCGAGTGGCTAATTTGAAACCAA  
 GAATCTCCTTTAATTTCTTTGTCACTGCACCTAAAAATGTGTCTGATATCATTCCTAGAACTGAAGTTGAAAAGG  
 CCATCAGGATGTCCCGGAGCCGTATCAATGATGCTTTCCGTCTGAATGACAACAGCCTAGAGTTTCTGGGGATAC  
 AGCCAACACTTGGACCTCCTAACCAGCCCCCTGTTTCCATATGGCTGATTGTTTTTGGAGTTGTGATGGGAGTGA  
 TAGTGGTTGGCATTGTATCCTGATCTTCACTGGGATCAGAGATCGGAAGAAGAAAAATAAAGCAAGAAGTGGAG  
 AAAATCCTTATGCCTCCATCGATATTAGCAAAGGAGAAAAATATCCAGGATTCCAAACACTGATGATGTTCAGA  
 CCTCCTTTTAAAAAATCTATGTTTTTCTCTTGGAGTATTTGTTGTATGTAATGTTAATTTTCATGGTATAG  
 AAAATATAAGATGATAAAGATATCATTAATGTCAAACCTATGACTCTGTTCAAGAAAAAATTTGTCCAAAGACA  
 ACATGGCCAAGGAGAGAGCATCTTCATTGACATTGCTTTTCAGTATTTATTTCTGTCTCTGGATTTGACTTCTGTT  
 CTGTTTCTTAATAAGGATTTTGTATTAGAGTATATTAGGGAAAGTGTGTATTTGGTCTCACAGGCTGTTTCAGGGA  
 TAATCTAAATGTAATGTCTGTTGAATTTCTGAAGTTGAAAACAAGGATATATCATTTGGAGCAAGTGTGGATCT  
 TGTATGGAATATGGATGGATCACTTGTAAAGGACAGTGCCTGGGAACCTGGTGTAGCTGCAAGGATTGAGAATGGCA  
 TGCATTAGCTCACTTTTCAATTAATCCATTGTCAAGGATGACATGCTTTCTTCACAGTAACCTCAGTTCAAGTACTA  
 TGGTGATTTTGCCTACAGTATGTTTGAATCGATCATGCTTTCTTCAAGGTGACAGGTCTAAAGAGAGAAGAATC  
 CAGGGAACAGGTAGAGGACATTGCTTTTTCTCACTTCCAAGGTGCTTGATCAACATCTCCCTGACAACACAAACTA  
 GAGCCAGGGGCCCTCCGTGAACCTCCCAAGAGCATGCCTGATAGAACTCATTTCTACTGTTCTCTAACTGTGGAGT  
 GAATGGAAATTCCACTGTATGTTCAACCTCTGAAGTGGGTACCCAGTCTCTTAAATCTTTTGTATTTGCTCACA  
 GTGTTTGAGCAGTGCTGAGCACAAGCAGACACTCAATAAATGCTAGATTTACAAAA

**FIGURE 72**

MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMN  
 NAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTILNTMSTI  
 YSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVVGKQLRPLYEEYVVLKNE  
 MARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAY  
 PSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFQKPNIDVTDAMVDQAWDAQRIKFKEAEKFF  
 VSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTAHHEMGH  
 IQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLK  
 QALTIVGTLPTTYMLEKWRWMVFKGEIPKDQWMKKWEMKREIVGVVEPVPHDETYCDPASLF  
 HVSDDYSFIRYYTRTLYQFQFQEALCQAAKHEGPHLHKCDISNSTEAGQKLL

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**N-glycosylation sites.**

amino acids 53-57, 90-94, 103-107, 322-326, 432-438, 546-550

**N-myristoylation sites.**

amino acids 260-266, 286-292, 395-401

**Cell attachment sequence.**

amino acids 204-207

**Neutral zinc metallopeptidases, zinc-binding region signature.**

amino acids 371-381

**FIGURE 73**

CCACGCGTCCGAGCGGGGTGGACAAGTGGCGTGTGTGCTGCGACCCCCGAGGGGAAGATGAACG  
GGACGCGGAACCTGGTGTACCCTGGTGGACGTGCACCCAGAGGACCAGGCGGCGGCGGGCAGGA  
AGACCTATGCCATGGTGTCCAGCCACTCAGCTGGTCATTCTCTGGCTTCAGAACTGGTGGAGT  
CCCATGATGGACATGAGGAGATCATTAAGGTGTACTTGAAGGGGAGGTCTGGAGACAAGATGA  
TTCACGAGAAGAATATTAACCAGCTGAAGAGTGAGGTCCAGTACATCCAGGAGGCCAGGAACT  
GCCTACAGAAGCTCCGGGAGGATATAAGTAGCAAGCTTGACAGGAACCTAGGAGATTCTCTCC  
ATCGACAGGAGATACAGGTGGTGTAGAAAAGCCAAATGGCTTTAGTCAGAGTCCCACAGCCC  
TGTACAGCAGCCACCTGAGGTGGACACCTGTATAAATGAGGATGTTGAGAGCTTGAGGAAGA  
CGGTGCAGGACTTGCTGGCCAAGCTTCAGGAGGCCAAGCGGCAACACCAGTCAGACTGTGTGG  
CTTTTGAGGTCACACTCAGCCGGTACCAGAGGGAAGCAGAACAAAGTAATGTGGCCCTTCAGA  
GAGAGGAGGACAGATGTCCAGAGTGAATTGGAGAATGTCCTGGGGGAATGAAGTTCCTTCCACA  
AACACAGCTCAGTTCCTTAGCAACAACTGTTTGTTTTTCTACTTGCTCCATCTGCAGCCTACG  
CTGCCCTGGCCTCCTGCAGACAGATAGTGGGGTTACCTGGCAAGGCCTGGTGAGAGCCAGTGA  
ACCTAAGCTTTGACTGGGTGGCCTTGTCTTTCTGGGGAGGAGGGAATGTACATTCAGGGAGTA  
GCCTTTTTCGGGAAAAATTCTCTAGGGCTACAGACAGTCATGTGTGACTTCTCTCTGCTGTGAA  
AACTCCCAGAGTCTCTTTAGGGATTTTCCCTAAGGTGTACCACCAGGCACACCTCAGTCTTCT  
TGACCCAGAGCCTGAAAACCTGTTTTACTGGGTTCCACCAGTCCCAGCAAAATCCTCTTTGTA  
TTTATTTTGCTAAGTTATTGGTGGTTTTGCTTACATCTCATGATTGATATAATACCAAAGTTC  
TATAGCCTTCTCTTGACGTATTTGGATTTGCTTGAAACCGGGAAAACCTGTTCCCATTAGGCTT  
GTTAATGTCAGAGTGACACTATTATGAATCTTTCTCTCCCTTTCCTCTGCCTGTTTCTTCTCT  
CTTCTCTCCTTCAAACCTTGCTCTGCAGCTAAGGAAGGTGAGTCTACTTTCCTGAGGCTTTGGG  
GTCAGAGTATATGTTGTTTGGAGAAAGAGGGCAATCAGGACTCTTCTGGGACCCAGATGAGTT  
CTTCACTAGCCCTTCTGAACCCCTTGCTCCATAATTGGTCTTTTATCCTGGCTCTGAATGACC  
CTGCAGGTCATCATGGTTTTCTTTTTTTATTGTTTTTTTTTTTTTCTGAGACAGAGTCTCACT  
CTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCAGCTCACTGCAACCTCTGCCTCCCGGAT  
TTAAGCGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGTGCCACCACGCCTG  
GCTGATTTTTGTATTTTATGATAGAGATGGGGTTTACCATACTGGCTAGGCTGGTCTCGAATT  
CCTGACCTCAGGTGATCCACCCACCTCGGCTTCCCAAAGTGCTAGGATTATAGGCTTGAGCTA  
CTGCGCCCCGGCCCATGGTGTTTTTCTTTAGGGCTCTTCCCTACAGCCTTGAGAAGTAGATAGGC  
ATCAGAGTATGGTACTATAGGAATCAGAAAAATTCAAAACAAATGTGGATTAAAGTGTTTAGGC  
TCTATGTGGCTCACGCAGCCAGAATCCTTAAGTCTGTGTGTTTCTGTGTCTCAAGACTGGGCT  
CACATTCTGGCTTTGTCCATAACAATGCTCTGGGATTTTCAGGGAGTTCCTTCATTTGTAAAAT  
GAGGGGGTCAGAGCAGGTGATATCCATGTTTCTTCCCTTTCTGATATTGTTGTCTGTGGCATA  
TTCTTTGTATGGCGAATTTAATAAATTATATTAATGTGTCA

## **FIGURE 74**

MNGTRNWCTLVDVHPEDQAAAGRKTYAMVSSHSAGHSLASELVESHGHEEIIKVYLKGRSGD  
KMIHEKNINQLKSEVQYIQEARNCLQKLREDISSKLDRNLGDSLHRQEIQVVLEKPNGFSQSP  
TALYSSPPEVDTCINEDVESLRKTVQDLLAKLQEAKRQHQSDCVAFEVTLSTRYQREAEQSNVA  
LQREEDRCPE

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-39

#### **N-glycosylation site.**

amino acids 2-6

#### **Amidation site.**

amino acids 21-25

GCTTGCACACATGGCTCCGGAGGCTCCGGTTGCCATCCGAGCCCTGCCAGGCTCTAACGTTCCCAACTGACAA  
CACCAGTAACTAAATATAGGAGCAGATGGTGGGGACGGGCTGTGCGAGCGGCTCCTTTGCAGAGGTCTCCGGACT  
GCAGATAAAGGCTCAGGCCCTTTTGTGAGAAGCAGACCAGCCTGGGGGCTGGCGGCAGGACACCTGTGTCTGC**ATG**  
CTGAAGAAGATGGGTGAGGCCGTGGCCAGAGTAGCAAGGAAGGTCAACGAGACGGTGGAGAGCGGCTCTGACACT  
CTGGACCTGGCGGAGTGCAAGCTGGTCTCCCTTTCCCATTTGGCATCTACAAGTCTCTGCGGAATGTCTCTGGCCAG  
ATCCACCTCATCACCTTGGCTAAACAACGACCTTAAGTCCCTCACCAGCAAGTTCATGACCACATTCAGTCAAGTCT  
CGAGAGCTCCACCTGGAGGGGAACTTCCTACCCGCTCCCCAGCGAGGTCAGTGCCCTGCAGCACCCTCAAGGCC  
ATTGACCTGTCCCGBAACAGTTCAGGACTTCCCTGAGCAGCTTACCGCCCTGCCGGCGCTGGAGACCATCAAC  
CTGGAGGAGAACGAGATCGTAGATGTGCCCGTGGAGAAGCTGGCCGCCATGCCAGCCTTGCGCAGCATCAACCTC  
CGCTTCAACCCACTCAACGCCGAGGTGCGCGTGATCGCCCCGCCGCTCATCAAGTTTGACATGCTCATGTCTCCG  
GAAGGCGCAAGAGCCCCCTACCT**TAG**GCCACCTCCTCATGCCACCCAGCAAGGGACAGAGGCCACAGGCCCTG  
GAACCTTGAAGGGAGGGAGGCCCATGGGAGGCCAAGCCTGGGGGCTGGGGGCGGGTGGGCCAGGAGCAGCAGTGG  
TGGGTGGGTGTCAGCTGGTCTGGATAGATAGCTTACAGCAGTAGTGGGCTCTGGAATGCCAAGGGAAGAGGCAA  
GGTGGGGCCTGCAGCCTGGACTCGGCACTCAGAGTCTGTGCAAACTCAGGCAGATCTCCTGCCCTCTGTAGC  
CTTGTCACTTGAAAAAACAGGACCTTTCCCTCCTTTGGGCTCCCTGGAGGTTTTTAAGCAGTACGTGCCCTCCA  
AGTTACCTCCAGATCAGCAGGCACAGGTGGGCATTGCCAGGTATTTTCTGAGCCCCCTGCGGGTTTTGAGGCCCTTGT  
TTTTAGTGCTGAGAGCCAGTTGCTGCCCTGAGAAGAGAAGACAACCTCCATCTATTTATTTGCTTCCCTGAGAAGT  
ACCTGGATGCGGCCCTCTGCAGGGCCCCAGTCTTCAGTCCCTGTGGTCCCTGGACTGGTGGGAACCTGAACTAGGAG  
TCCTGGGAGAGCTGTGGTGGGAATATGGGCTGGCACTCTGCAGGGCAAGAACATTCATGTAGGAGCCCCGAGGAC  
CANCANGCTGGGAATGGGGAGCAAGTCAGTCAAGTCTGTCTATTTCCCAAGTTAAACAAATTTGGCGGGGTGGGAA  
GTCTTGAGTGCTCCGTCCCTCTAGCATCACTCCTGAGCTGCGGGAGAGGTGGCCCAGAGAACAGCAGAGTCAAGTT  
ACACCTGCAGCTCTTGCTCTAAAGTGATTAGATGGCCACCCTCACCAGTGTCCAGTCCAGCAGCAGCCTGGCTGCC  
TTGTCTAGGCCCTCCTGGGGGCGAGAAGGCGATGTGGACCACGGGATTTGTAGCCAGCCAGCTCCCAGGCCAACGCC  
CAAAGCCCTGATGACCTGGTCTCTTCTGAGGCCCTCAACCTGGCATCTTAGGGTATGGTCAGGCAACAGGGTGACC  
AGCTGTCTGGTTTTCCAGGACATGGAACCTTCAATGCTAAAACTGGGACATTTACCAGCAAGTGGGGATGGTTG  
GTCCCCTACCAGGAGAGGGGCTGGGGCTTTGCTTCCCGAAGACCGCTGTGGCTTGAGAACCTTGACTGCTTGG  
TCCTCAGGTATCTACCTCCACCTTCTCCTCATCTGTGGAGCAAGCCAACTCAGTGCCCCAGACCCCCACCTGATC  
TGCATCTTTGTTTGCTCCAGAGACACCTGAGGCCCCAGAGCTTGAGGCAAAGCCAGGCCGCTCCAAATCCTGTGTG  
CCGTGGACGAGTGGCCACTTTACTACTCCTAAGGCTAAGATGTTTGAGAGTCTAGACCCTGCTCAGAGCAGTAAT  
CCCTGCTCAGAATGCTCCCAAGTTCCTCGTCCCTGCCCAGGTCCTTGTCTCTTGGGAAGGAACTGATAGGTCGG  
GCCATTGTTGGGCCATCACTGAGCGCTCAGTATCTCAAGAGCTCTGTTCACTTCTGCTATCCCCAAGGCCCTGG  
TTGGTCAAACCTCTGGGCAAAGGGTTTTTCAGGATGAGGAGTCAAGACAGGATGTCACAGAGTACCAGGTTCACTC  
GTGGGTGTTGGGGCAAGTGGGGCTGAAGTCTGTGCAGGCTGCGCTGGCCCCACCTGCCCTTGTGCCCTGGAGT  
GGGGTTTTCTCCTTGTGTAAGAAGAGGCATCCTTCTCTGATGTGCACAAACACAATGTATGACCAGAGCCTTGCAA  
CTCAAAGTGTGGTCTGTGGACCAGCAGCGGCAGTGACACCTGGGAGCTTGTTAGGGAATGCAGAGTCTAGGCCTCA  
CCCTATACCTCCCGACTCAGACCTCGACTTTTAGCAAGACCCCCAGCTGATTCTATAAGCACTTTAGAGTTTGA  
AAGCAAGGACCTAGGCTGGGGATGTCTCCGACGAGAGGTTGAAGTTCTCATGTTCTCTCCCTGCCACTTCC  
AGGGATCTGAGCCTGTGTTTCAAGCTCCTCCCTAACCACCCCTGGGAGACACTTGGCTGTAGATTGTCCAGTCC  
TCTGCATGGCACTCCTGAAGAAGGGAGTGTGACCTGCAGTCACCAGGAGATGAGGGTTAGGTGTGCCCAGCCCTC  
CAGACCCGGCCTTTCTGTTAACCCTGCATGCCAAGCTGCCTGCTGCCCCAGGTCCTCACTCAGGCCTTTGAA  
GGGCGAGCTTCTGGAAAGTTGTTTTCTCCTCTGCTGGAGAGCTTTGCCCTTGTCTGTCTTGAAAGTGTGGGCAGC  
ACAGATGCCCTCCAAATCAGAGCTCAGAGTGAGTGGCCCTTAAGCTTCAGTCTGCAATAAAGAATGCATTGGTT  
TCAA

**FIGURE 76**

MLKKMGAVARVARKVNETVESGSDTLDLAECKLVSFPIGIYKVLNRVSGQIHLITLANNELK  
SLTSKFMTTFSQLRELHLEGNFLHRLPSEVSALQHLKAIDLSRNQFQDFPEQLTALPALETIN  
LEENEIVDVPVEKLAAMPALRSINLRFNPLNAEVRVIAPPLIKFDMLMSPEGARAPLP

Important features of the protein:

N-glycosylation sites.

amino acids 17-21, 47-51





## **FIGURE 78**

MDDDLMLALRLQEEWNLQEAERDHAQESLSLVDASWELVDPTPDLQALFVQFNDQFFWGQLEA  
VEVKWSVRMTLCAGICSYEGKGGMC S IRLSEPLLKLRPRKDLVETLLHEMIHAYLFVTNNDKD  
REGHGPEFCKHMRINSLTGANITVYHTFHDEVDEYRRHWWRCNGPCQHRPPYYGYVKRATNR  
EPSAHDYWWAEHQKTCGGTYIKIKEPENYSKKGKGKAKLGKEPVLAAENKGTFFVYILLIFM

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-41

#### **N-glycosylation sites.**

amino acids 148-151, 217-220

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 184-187

#### **Casein kinase II phosphorylation sites.**

amino acids 30-33, 121-124, 154-157, 187-190, 192-195

#### **Tyrosine kinase phosphorylation site.**

amino acids 211-218

#### **N-myristoylation sites.**

amino acids 59-64, 85-90, 146-151

#### **Neutral zinc metallopeptidases, zinc-binding region signature.**

amino acids 108-117

**FIGURE 79**

CGGACGCGTGGGTGGCAACCAGGAGAAGCCAAACTTGGTCCCCGGCTCGCGGAGTGCCTGCG  
AGCGGTGCTC**ATG**GCGCTCTATGAGGTCTTCTCTCACCCGGTCGAGCGCAGTTACCGCGCGGG  
GCTCTGCTCCAAAGCCGCGCTGTTCTGCTGCTGGCCGCTGCGCTCACGTACATCCCGCCGCT  
GCTGGTGGCCTTCCGGAGCCACGGGTTTTGGCTGAAGCGGAGCAGCTACGAGGAGCAGCCGAC  
CGTGCGCTTCCAACACCAGGTGCTGCTCGTGGCCCTGCTCGGACCCGAAAGCGACGGGTTCCT  
CGCCTGGAGCACGTTCCCCGCCTTCAACCGGCTGCAAGGGGATCGCCTGCGCGTCCCGCTCGT  
TTCGACTAGAGAAGAAGACAGGAACCAGGATGGGAAGACGGACATGTTACATTTTAAGCTGGA  
GCTTCCCCTGCAGTCCACGGAGCACGTTCTCGGTGTGCAGCTCATCCTGACTTTCTCCTATCC  
ATTACACAGGATGGCGACCCCTCGTGATGCAGAGCATGGCGTTTCTCCAGTCCTCCTTTCTGT  
CCCGGGATCCAGTTATACGTGAACGGAGACCTGAGGCTGCAGCAGAAGCAGCCGCTGAGCTG  
TGGTGGCCTAGATGCCCGATACAACATATCCGTGATCAACGGGACCAGCCCCTTTGCCTATGA  
CTACGACCTCACCCATATTGTTGCTGCCTACCAGGAGAGGAACGTTACCACCGTCCTGAATGA  
TCCCAACCCCATCTGGCTGGTGGGCAGGGCCGCAGATGCTCCATTTGTGATTAATGCTATCAT  
CCGATACCCTGTGGAAGTCATTTCTTATCAGCCAGGATTCTGGGAGATGGTAAAGTTCGCCTG  
GGTACAGTATGTGAGCATCCTGCTTATCTTCCTCTGGGTGTTTGAAAGAATCAAGATCTTCGT  
GTTTCAGAATCAGGTGGTGACCACCATTCCTGTGACAGTGACGCCCCGGGGAGACTTGTGTAA  
GGAGCACTTATCC**TAG**AAGGCCATTTCTGAAGACTCAGCAGGACCGTGGCTGCCTCATTGTC  
ATCTTCTGGGAACATCTTAGGACCTTTTGAAAGAGCCCAGCGGACACCTGCGGGCTTGTGTGC  
TTTTCCCTCAGAGACAACGGTTCTTCCGGTTTTGCTCTACACAGTTCCGTATCTTCAGAGCT  
CCTGCAGAATTGTCAGGGACTAGTTTGTGGAAAGGTCTGAGAGTTCTGGAGGCTATAATTAG  
CTTTTTGGGTTTTCTTCTTGCCTTAGCGTTGAATTTCAGGAGAAAATTGCAGTCAGTTAG  
ACATCTTGAAAGAGTCCCATCTCTGGTCAAGCAGAGACTTTTCTCTGTTGAAGTGAAGAAC  
ACACTGTGCATTTCTTCCTTCTGTTGTGAGCCACTCTTACTCTTTTCAGGGCTCTCTTGTGAC  
AAACATGCCAATCACTAGCACTTTGCACCCCTGGGCTTCTCCATTTCCATTACACAGCTTTGA  
TTTCCAGAGCTGAGGCCTTTAACTGGAGACCTGGAGGGGCAGGGCCCAAGGGCAAGGGCCGCA  
TTAGCACAGGCAATCAGGGAGGGCCGCTGAAGGACACTTGGACCGTCCACCTGCCCCAGCCCA  
ACAGTCAGTCATCTGTCATCAGCTCAGCTGAGCAGCCCTGGATCTTTGCCGTACTGTGACTGG  
GCTCTTTGCCCTATTTTTTCCCTCTGTCTGTGCCCTGGATGGCAGGCTGAAGTCAGAGGGGCT  
GTTTCATTCTCAGCCCCCTCAGCAGCACTGGGGGAAGAAAGCATTGTCACAACAGGTTCTTTC  
TGGCCCTCACCCAACAGCCTGGGCACTTGGCCCTCCTCCTCCTTGACAGCCCTCCCCCTTCCT  
GCAAAGGACAGGGGCGACAGGGGTTGGTGTGGGATTGGCTCCCGCTGCCTGACAACCACAAG  
TTTATTTGGAAGGCTAGCGGGAAGCCCAGCGGCTGGCGTTTCCCTTGACTAAGGAACAGGGTG  
CCCATCAGAGTGGGGCGGGCAGCTTTGGGAAGGACACAAGAAGCAGTAAGAGTGTAAGAGAGGA  
TGCTGGCCTGGGCAGGCCAGTCCAGCCTGGCCACTAGCAGAATACCAAGCAGTCCAGTGGATT  
ACCCTCGTGGCTAAGCAAGTGTCTGCAGGAGCAGAGATGGCTGGAAGGGGGCTCTGCACACGG  
AAGATGGCTTGTTACAGCCATTACCTCCTGAGGATGTGGGCAGTCTCCTCCAAGAACACATG  
GAGCTGCTTCCCTGATCCCAAGCAGGTCATTGCCACTGGAAGGACATGGCCCCGGTGATCCATG  
CTTCATGCCCAACCCAGAAACACACCCCTCAGTGTGTGCCTCAGTTTACTTTGGAGATCAGTTG  
TCGTTTTTTAGTGCTCCTTTAGGCTTACTAAAACAGTTTTTGGAACAAAGCTATTTTGAAGTAT  
TCAAGCAGAGGAATTCCTAACACTGACCCCTTGTCTTTTTTTTAAATATTCAGGCTGTTTTAT  
ATGCCATAAATTTTTTTCTTAAGATCTAAACGAAAAATAGTTTCTTGTTTAAATTCACATAAG  
CAATGAGATATGGAAAGATGACAAGATACGTATAAACATTGGTTTTGCATCTTATTAAATTAT  
CTAATGCAATCTTGTATAAAGAACCCATGATGTTTTGTAACTTTCTAATTAAATGTTCAA  
ATGAG

## **FIGURE 80**

MALYEVFSHPVERSYRAGLCSKAALFLLLAALTYIPLLVAFRSHGFWLKRSSYEEQPTVRF  
QHQVLLVALLGPESDGFLAWSTFPAFNRLQGDRLRVPLVSTREEDRNQDGKTDMLHFKLELPL  
QSTEHLVGLVQLILTFSYRLHRMATLVMQSMFLQSSFPVPGSQLYVNGDLRLQQKQPLSCGGL  
DARYNISVINGTSPFAYDYDLTHIVAAYQERNVTTVLNDPNPIWLVGRAADAPFVINAIIRYP  
VEVISYQPGFWEMVKFAWVQYVSILLIFLWVFERIKIFVFQNVVTTIPVTVTPRGDLCKEHL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-34

#### **Transmembrane domain:**

amino acids 268-284

#### **N-glycosylation sites.**

amino acids 194-198, 199-203, 221-225

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 51-55

#### **Tyrosine kinase phosphorylation site.**

amino acids 250-259

#### **N-myristoylation site.**

amino acids 187-193

#### **Cell attachment sequence.**

amino acids 307-310

GCCGGGAGCTTCCCTG**GATG**GTGCCCGCCGCTCCGAGCCGGGGAGGAGCTGCCAGGGGGCCAGCTGGGCAGGAGCCT  
 GGGTCCGCTGCTGCTGCTCCTGGCGTTGGGACACACGTGGACCTACAGAGAGGAGCCGGAGGACGGCGACAGAGA  
 AATCTGCTCAGAGAGCAAAATCGCGACGACTAAATACCCGTGTCTGAAGTCTTCAGGCGAGCTCACCACATGCTA  
 CAGGAAAAAGTGCTGCAAAGGATATAAATTTGTCTTGGACAATGCATCCCAGAAGATTACGACGTTTGTGCCGA  
 GGCTCCCTGTGAACAGCAGTGCACGGACAACATTTGGCCGAGTGCTGTGTACTTGTATTCCGGGATACCCGATATGA  
 CCGGGAGACACCGGAAGCGGGGAGAAGCCATACTGTCTGGATATTGATGAGTGTGCCAGCAATGGGACGCT  
 GTGTGCCACATCTGCATCAATACCTTTGGGACGCTACCGCTGCGAGTGCCGGGAAGGCTACATCCGGGAAGATGA  
 TGGGAAGACATGTACCAGGGGAGACAAATATCCCAATGACACTGGCCATGAGAAGTCTGAGAACATGGTGAAAG  
 CGGAACCTTGCTGTGCCACATGCAAGGAGTTCTACCAGATGAAGCAGACCGTGCTGCAGCTGAAGCAAAAGATTGC  
 TCTGCTCCCCAACAAATGCAGCTGACCTGGGCAAGTATATCACTGGTGACAAGGTGCTGGCCTCAAACACCTACCT  
 TCCAGGACCTCCTGGCCTGCCCTGGGGGCCAGGGCCCTCCCGGCTCACCAGGACCAAAGGGAAGCCAGGCTTCCC  
 CGGTATGCCAGGCCCTCCTGGGCAGCCCGGCCACGGGGCTCAATGGGACCCATGGGACCATCTCCTGATCTGTC  
 CCACATTAAGCAAGGCGGAGGGGCCCTGTGGGTCCACAGGGGCACAGGAAGAGATGGTTCTAAGGGGAGAG  
 AGGAGCGCCTGGGGCCAGAGGGTCTCCAGGACCCCTGGTTCTTTCGACTTCTCTGCTACTTTATGCTGGCTGACAT  
 CCGCAATGACATCACTGAGCTGCAGGAAAAGGTGTTGGGGACCGGACTCACTCTCAGCAGAGGAGTTCCCTTT  
 ACCTCAGGAATTTCCAGCTACCCAGAAGCCATGGACCTGGGCTCTGGAGATGACCATCCAAGAAGAACTGAGAC  
 AAGAGACTTGAGAGCCCCCAGAGACTTCTACCC**TAG**CACATCCCAACACCGTCACGCCAAAGGAAGAGAAAGAT  
 CACTCACCTGCAGTTAAACCATCTAAAGAGAAGAAAGACCACCTGGAGACCTAGAAAACATACATATTTTCTCTCT  
 TCTTCTCCTGACGTCTCTCCACTCCTCTTCTTCCAAATACGATGCTATTTTCAGAGTCCCCTCCTAGGGCTCGAG  
 ACATGAGGGAGTGAATGATTGATTACCTGCTTCTCTCAATAGAGTCCATTGGGGTGGTTTGCATTGTAACTTTTCT  
 TTTTACATCCTATTTTTCCAGGAACCTTTGGATTAAAGTACTCTCACAGTGTCTTAAATCATAAAATTTCTGAAGTT  
 AAATTTGGCAGAGTATCAAAAGGGGGAAAATGACAAAGTGAGCTCTAAGAAAATGTGAGGCTACTTCTAAGATGT  
 GTGTTACAAATAGACCATAACTCCTCTAGTATCAAAATTTGGGGCTCTTCAGTTAAAAAGGGGTGGGGAGGACAAA  
 CGTGTCGATGTGCTTTGGTGGAGAATTTTTCTTGTGCTTCTAGTAGACTTTAAATATTGTATCCCTTTGTCAA  
 ACCTTGTTTCCCAATTCATTAAGAGAGGAGAGAATTAATGGCGTTTAGAGAAGATAGAAAAGATCACAGT  
 CTATATATTACTGTTATATAGATTGCCACATTTCTAAATCTCAAAATACGGTGCTTAAGGTTTCATGCCATGCTTAT  
 CTGTAAGTATCCTATTTAGGGAAGAAGATTAAACTCTCTTTTCAAAAAAACAAAGTGAAATGCCTGGATTACAT  
 TAAAACAATGGGCTCTCGTTTGCTATAATATTTTAAAGCTGTTTAAATCAACAGTGGAGTCTGCTCTATAAATATA  
 GATTATTTGTTCATAAACTGGCTGAGCTTAGAGAGAGGTGCAGAATTCCTGGTCTGAGCAGGTGCCAGAAGG  
 TACCATTAGGTGCCATGATCCAGGCTGAACCAATATACAGTGGGGCTGAAGTCTGCAAGGAGGTTGCTGGCTTGG  
 GCTGACCTCACTAATGCCATCAGCAGCGGTAGGTAAATTTTTCTCCTTGGGTATTACAAGTTTTTGTCTGGAGC  
 CAACCAAGCTTGCCACCAACATATTGAGGATAATACACTATTTGAAAGTTATCTTGGATGGGGAGAAAAAATA  
 GTGGTTTTCTTGTGTGCAAAAACTTCTTCTTCTATTTCTCATTTTTTCTTAAATTTTCTTAAATTTAGTCCAAGTTC  
 CAGTTCCTTTTAGGCCTTCTCTTTGATTTATTTTCCCCTGCATGTGAGAAGCAGTTCAGAAAAAGGTCTATATCTC  
 CACCTCCTAGTGAGTTAGAGTGTTTTCTCAGAGCACCTCTGGGTGGCAAAGGGAAGCATGTTTCTGCCAAGGTTT  
 GCTGTGGATTGAGAAGCACCCAGGAGCAAGAGACCAGAAGGATGATCTGCTCCTTTGTAACGTTGTTGAGGGCCCT  
 CTTGTTTCCAATGAGCAGCTTATAGTTTACTCAGAGTCCACTTTCTCACTGGACACAAAGTGGCTCTTTATCT  
 ACCTTTGCGGGAGATTTTCACTCTCCTGTCAAATGATCGTTCTCACACTCATATTAGCTCATGTTGGAAATTTCCCA  
 TCCTGCCATGTCTTTCCCATTTCTTTTTGGCTTTTTTGGCTCCACCTTTTAGCCACATCATTTAACTCCACTA  
 CTGTGAAAGCTTGCTTAAAGAAAATCCCTCTTGCCGGGTGTGGTAGCCACGCCTCTAATCCCAGCACTTTGGG  
 AGGCTGAGGCGGGGAGATCACAAGGTCAGGAGATCGAGACCAGCCTGACCAACATGGTGAAACCTGTCTCTACT  
 AAAAAATACAAAAATTAGCTGGGCGTGTTGGCACACACCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAA  
 TTACTTTAACTGCGGGGGGAGCCTAGATTGCGTACTGCACTCCAGCCTAGGCAACAGAGGGAGACTCTGTCTC  
 ATTAATAA

**FIGURE 82**

MVPPPPSRGGAARGQLGRSLGPLLLLLLALGHTWTYREEPEDGDREICSESKIATTKYPCLKSS  
 GELTTCYRKKCKGYKFVLGQCIPEDYDVCAEAPCEQQCTDNFGRVLCTCYPGYRYDRERHRK  
 REKPYCLDIDECASSNGTLCAHICINTLGSYRCECREGYIREDDGKTCTRGDKYPNDTGHEKS  
 ENMVKAGTCCATCKEFYQMKQTVLQLKQKIALLPNNAADLGKYITGDKVLASNTYLPGPPGLP  
 GGQGPSPGSPGKSPGFPGMPGPPGQPGPRGSMGPMGSPDLSHIKQGRRGPVGPPGAPGRDG  
 SKGERGAPGPRGSPGPPGSFDFLLMLADIRNDITELQEKVFGHRTHSSAEFFPLPQEFPSYP  
 EAMDLSGDDHPRRTETRDLRAPRDFYP

**Important features of the protein:****Signal peptide:**

amino acids 1-34

**N-glycosylation sites.**

amino acids 142-148, 182-188

**Tyrosine kinase phosphorylation site.**

amino acids 125-132

**N-myristoylation sites.**

amino acids 10-16, 143-149, 155-161, 196-202, 250-256

**Amidation site.**

amino acids 299-303

**Aspartic acid and asparagine hydroxylation site.**

amino acids 150-162

**Cell attachment sequence.**

amino acids 176-179

**Clq domain proteins.**

amino acids 247-280

**Calcium-binding EGF-like domain proteins pattern proteins.**

amino acids 144-165

**FIGURE 83**

ATCTGAGTGAGCTAACTGACACA**ATG**AAACTGTCAGGCATGTTTCTGCTCCTCTCTCTGGCTC  
TTTTCTGCTTTTTTAACAGGTGTCTTCAGTCAGGGAGGACAGGTTGACTGTGGTGAGTTCCAGG  
ACCCCAAGGTCTACTGCACTCGGGAATCTAACCCACACTGTGGCTCTGATGGCCAGACATATG  
GCAATAAATGTGCCTTCTGTAAGGCCATAGTGAAAAGTGGTGGAAAGATTAGCCTAAAGCATC  
CTGGAAAATGCT**TGA**GTAAAGCCAATGTTTCTTGGTGACTTGCCAGCTTTTGCAGCCTTCTTT  
TCTCACTTCTGCTTATACTTTTGCTGGTGGATTCTTTAATTCATAAAGACATACCTACTCTG  
CCTGGGTCTTGAGGAGTTCAATGTATGTCTATTTCTCTTGATTCACTTGTCAATAAAGTACATTC  
TGCAAAAGCAAAAA

## **FIGURE 84**

MKLSGMFLLLSLALFCFLTGVFSQGGQVDCGEFQDPKVYCTRESNPHCGSDGQTYGNKCAFCK  
AIVKSGGKISLKHPGKC

**Important features of the protein:**

**Signal peptide:**

amino acids 1-23

**N-myristoylation sites.**

amino acids 26-32, 52-58, 56-62, 69-75

**Kazal serine protease inhibitors family signature.**

amino acids 40-63





## **FIGURE 86**

MQAPGRGPLGLRLMMPGRRGALREPGGCGSCLGVALALLLLLLPACCPVRAQNDTEPIVLEGK  
CLVVCDSSPSADGAVTSSLGISVRSKSAKVAFSATRSTNHEPSEMSNRTMTIYFDQVLNIGN  
HFDLASSIFVAPRKGIIYSFSFHVVVKVYNRQTIQVSLMQNGYPVISAFAGDQDVTREAASNGVL  
LLMEREDKVHLKLERGNLMGGWKYSTFSGFLVFPL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-48

#### **N-glycosylation sites.**

amino acids 53-57, 110-114

#### **N-myristoylation sites.**

amino acids 26-32, 27-33, 29-35, 33-39, 76-82, 205-211

#### **Amidation site.**

amino acids 16-20

#### **C1q domain signature.**

amino acids 117-148

#### **C1q domain proteins.**

amino acids 115-149



**FIGURE 88**

MAAASRSASGWALLLLVALWQQRAAGSGVFQLQLQEFINERGVLASGRPCPEPGCRTFFRVCLK  
 HFQAVVSPGPCTFGTVSTPVLGTNSFAVRDDSSGGGRNPLQLPFNFTWPGTFSLIIIEAWHAPG  
 DDLRPEALPPDALISKIAIQGSLAVGQNWLLDEQTSTLTRLRYSYRVICSDNYYGDNCSRLCK  
 KRNDHFGHYVCQPDGNLSCLPGWTGEYCQOPICLSGCHEQNGYCSKPAECLCRPGWQGRLCNE  
 CIPHNGCRHGTCSTPWQCTCDEGWGGLFCDQDLNYCTHHS PCKNGATCSNSGQRSYTCTCRPG  
 YTGVDCELELSECDSNPCRNGGSKDQEDGYHCLCPPGYGLHCEHSTLSCADSPCFNGGSCR  
 ERNQGANYACECPPNFTGSNCEKKVDRCTSNPCANGGQCLNRGPSRMCRCRPGFTGT YCELHV  
 SDCARNPCAHGGTCHDLENGLMCTCPAGFSGRRCVVRTSIDACASSPCFNRATCYTDLSTDTF  
 VCNCYPYGFVGSRCFFPVGLPPSFPWVAVSLGVGLAVLLVLLGMVAVAVRQLRLRRPDDGSREA  
 MNNLSDFQKDNLI PAAQLKNTNQKKELEVDCGLDKSNCGKQQNHTLDYNLAPGPLGRGTM PGK  
 FPHSDKSLGEKAPLRLHSEKPECRISAICSPRDSMYQSVCLISEERNECVIATEV

**Important features of the protein:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 530-552

**N-glycosylation sites.**

amino acids 108-112, 183-187, 205-209, 393-397, 570-574, 610-614

**Glycosaminoglycan attachment site.**

amino acids 96-100

**Tyrosine kinase phosphorylation site.**

amino acids 340-347

**N-myristoylation sites.**

amino acids 42-48, 204-210, 258-264, 277-283, 297-303, 383-389,  
 415-421, 461-467, 522-528, 535-541, 563-569, 599-605, 625-631

**Amidation site.**

amino acids 471-475

**Aspartic acid and asparagine hydroxylation site.**

amino acids 339-351

**EGF-like domain cysteine pattern signature.**

amino acids 173-185, 206-218, 239-251, 270-282, 310-322, 348-360,  
 388-400, 426-438, 464-476, 506-518

**Calcium-binding EGF-like:**

amino acids 224-245, 255-276, 295-316, 333-354, 373-394, 411-432,  
 449-470



**FIGURE 90**

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENK  
PGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSG  
LDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTL  
EDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKVTMAAIQDGLAKGENDETVSNT  
LTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVK  
YGTISPEEGVSYLENLDEMIALQTKNKLEKNATDNISKLFPAPEKSHEETDSTKEEAAKMEK  
EYGS�KDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFIN  
KQADAYVEKGILDKEEAEAIKRIYSSL

**Important features:****N-glycosylation sites:**

amino acids 68-71, 346-349, 350-353

**Casein kinase II phosphorylation site:**

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-  
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-  
341, 369-372, 382-385, 386-389, 387-390

**N-myristoylation sites:**

amino acids 143-148, 239-244

**FIGURE 91**

TGCATCAGTGCCCAGGCAAGCCCAGGAGTTGACATTTCTCTGCCCAGCC**ATG**GGCCTCACCCCT  
 GCTCTTGCTGCTGCTCCTGGGACTAGAAGGTCAGGGCATAGTTGGCAGCCTCCCTGAGGTGCT  
 GCAGGCACCCGTGGGAAGCTCCATTCTGGTGCAGTGCCACTACAGGCTCCAGGATGTCAAAGC  
 TCAGAAGGTGTGGTGCCGGTTCTTGCCGGAGGGGTGCCAGCCCCTGGTGTCTCAGCTGTGGA  
 TCGCAGAGCTCCAGCGGGCAGGCGTACGTTTCTCACAGACCTGGGTGGGGGCCTGCTGCAGGT  
 GGAAATGGTTACCCTGCAGGAAGAGGATGCTGGCGAGTATGGCTGCATGGTGGATGGGGCCAG  
 GGGGCCCCAGATTTTGCACAGAGTCTCTCTGAACATACTGCCCCCAGAGGAAGAAGAAGAGAC  
 CCATAAGATTGGCAGTCTGGCTGAGAACGCATTCTCAGACCCTGCAGGCAGTGCCAACCCTTT  
 GGAACCCAGCCAGGATGAGAAGAGCATCCCCTTGATCTGGGGTGCTGTGCTCCTGGTAGGTCT  
 GCTGGTGGCAGCGGTGGTGCTGTTTGCTGTGATGGCCAAGAGGAAACAAGAATCCCTCCTCAG  
 TGGTCCACCACGTCAG**TGA**CTCTGGACCGGCTGCTGAATTGCCTTTGGATGTACCACACATTA  
 GGCTTGACTCACCACCTTCATTTGACAATACCACCTACACCAGCCTACCTCTTGATTCCCCAT  
 CAGGAAAACCTTCACTCCCAGCTCCATCCTCATTGCCCCCTCTACCTCCTAAGGTCCTGGTCT  
 GCTCCAAGCCTGTGACATATGCCACAGTAATCTTCCCGGGAGGGAACAAGGGTGGAGGGACCT  
 CGTGTGGGCCAGCCCAGAATCCACCTAACAAATCAGACTCCATCCAGCTAAGCTGCTCATCACA  
 CTTTAAACTCATGAGGACCATCCCTAGGGGTTCTGTGCATCCATCCAGCCAGCTCATGCCCTA  
 GGATCCTTAGGATATCTGAGCAACCAGGGACTTTAAGATCTAATCCAATGTCCTAACTTTACT  
 AGGGAAAGTGACGCTCAGACATGACTGAGATGTCTTGGGGAAGACCTCCCTGCACCCAACTCC  
 CCCACTGGTTCTTCTACCATTACACACTGGGCTAAATAAACCCTAATAATGATGTGCAAAAAA  
 AA

## **FIGURE 92**

MGLTLLLLLLLLGLEGGIVGSLPEVLQAPVGSSILVQCHYRLQDVKAQKVWCRFLPEGCQPLV  
SSAVDRRAPAGRRFTLTDLGGGLLQVEMVTLQEEDAGEYGC MVDGARGPQILHRVSLNILPPE  
EEEETHKIGSLAENAFSDPAGSANPLEPSQDEKSIPLIWGAVLLVGLLVAAVVLFAVMAKRKQ  
ESLLSGPPRQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-15

#### **Transmembrane domain:**

amino acids 161-181

#### **N-myristoylation sites.**

amino acids 17-23, 172-178

#### **Amidation site.**

amino acids 73-79



**FIGURE 93**

GGCGCGCTTGCCGGGCTCTCCGGAAGGAGACGTGGCGGCGGTTGGGCCGGTGATACCCGGGCG  
CTTTATAGTCCCGCCGCCTCCTCCTCCACCTCCTCCTCCTCCTCCTCCTCCTGCGGCGAGAG  
GAGGTTGTGGCGGTGGCTGGAGAAAGCGGCGGCGGAGGATGGAGGAAGGAGCGGCGGCGGTAC  
GGAGTCTGGTCCC GGCGGGCCGGTGTTACTGGTCTCTGCGGCCTCCTGGAGGCGTCCGGCG  
GCGGCCGAGCCCTTCTCAACTCAGCGATGACATCCCTTTCCGAGTCAACTGGCCCGGCACCG  
AGTTCTCTCTGCCCCAACCTGGAGTTTTATATAAAGAAGATAATTATGTCATCATGACAACTG  
CACATAAAGAAAAATATAAATGCATACTTCCCCTTGTGACAAGTGGGGATGAGGAAGAAGAAA  
AGGATTATAAAGGCCCTAATCCAAGAGAGCTTTTTGGAGCCACTATTTAAACAAAGCAGTTGTT  
CCTACAGAATTGAGTCTTATTGGACTTACGAAGTATGTCATGGAAAACACATTCGGCAGTACC  
ATGAAGAGAAAAGAAACTGGTCAGAAAATAAATATTACGAGTACTACCTTGGAATATGTTGG  
CCAAGAACCCTTCTATTTGAAAAAGAACGAGAAGCAGAAGAAAAGGAAAAATCAAATGAGATT  
CCACTAAAAATATCGAAGGTGAGATGACACCATACTATCCTGTGGGAATGGGAAATGGTACAC  
CTTGATGTTTTGAAACAGAACCGGCCAGATCAAGTACTGTGATGTACATATGTCATCCTGAAT  
CTAAGCATGAAATTCTTTTCAGTAGCTGAAGTTACAACCTTGTGAATATGAAGTTGTCATTTTTG  
CACCCTCTTGTGCAGTCACTCTAAATATAGGTTTCAGAGCATCTCCTGTGAATGACATATTTT  
GTCAATCACTGCCAGGATCTCCATTTAAGCCCCTCACCTGAGGCAGCTGGAGCAGCAGGAAG  
AAATACTAAGGGTGCCTTTTAGGAGAAATAAAGAGGGTGTGCGGTTGGTGGAAATATGAATTCT  
GCTATGGCAAACATGTACATCAATACCATGAGGACAAGGATAGTGGGAAAACCTCTGTGGTTG  
TCGGGACATGGAACCAAGAAGAGCATATTGAATGGGCTAAGAAGAATACTGCTAGAGCTTATC  
ATCTTCAAGACGATGGTACCCAGACAGTCAGGATGGTGTACATTTTTTATGGAAATGGAGATA  
TTTGTGATATAACTGACAAACCAAGACAGGTGACTGTAAAACTAAAGTGCAAAGAATCAGATT  
CACCTCATGCTGTTACTGTATATATGCTAGAGCCTCACTCCTGTCAATATATTCTTGGGGTTG  
AATCTCCAGTGATCTGTAAAACTCTTAGATACAGCAGATGAAAATGGACTTCTTTCTCTCCCCA  
CTTAAAGGATATTAAAGTTAGGGGAAAGAAAAGATCATTGAAAGTCATGATAATTTCTGTCC  
ACTGTGTCTCATTATAGAGTTCTCAGCCATTGGACCTCTTCTAAAGGATGGTATAAAATGACT  
CTCAACCCTTTGTGAATACATATGTGTATATAAGAGGTTATTGATAAACTTCTGAGGCAGAC  
ATTTGTCTCGCTTTTTTTTCATTTTTTGTGTTGTGCTTATAAACTGACTGTTTTTCTTTGCTTGG  
TACTGTGATTCCAAAATAAATCTCATCCAAGCAAGTTAGAGTCCAGCCTAATCAAATGTCATA  
ATTGTTGTACCTATTGAAAGTTTTTAAATAATAGATTTATTATGTAAATTATAGTATATGTAA  
GTAGCTAATGAAGTAAAGATCATGAAGAAAGAAATTGATAGGTGTAAATGAGAGACCATGTAA  
AATATGTAAATTCTAGTACCTGAAATCCTTTCAACAGATTTTTTATATAGCAACTGCTCTCTGC  
AAGTAGTTAAACTAGAACTGGGCACATGGTAGAGGCTCACATGGGAGTTGTCCTCACCCCTG  
TTAATCTCAAGAACTCTTATTTATAATAGGTTGCTTCTCTCTCAGAACTTTTATCTATTACT  
TTTTTCTTCTTATGAGTATGTTTACTCTCAGAGTATCTATCTGATGTAGACAGTTGGTGATGC  
TTCTGAGACTCAGAATGGTTTACTCTAACAAAACACTGTGCTGTCTATCCCTTGTAATTGCGCT  
ACTGTAATATGGATTTCACTTCTGAACAGTTTACAGCACAATATTTATTTTAAAGTGAATAAA  
ATGTCCACAAGCAAAA

**FIGURE 94**

MEEGGGGVRS LVPGGPVLLVLCGLLEASGGGRALPQLSDDIPFRVNWPGTEFSLPTTGVLYKE  
DNYVIMTTAHKEKYKCILPLVTSGDEEEEEKDYKGPNPRELLEPLFKQSSCSYRIESYWTYEVC  
HGKHIRQYHEEKETGQKINIHEYLLGNMLAKNLLFEKEREAEKEKSNEIPTKNIEGQMTPLY  
PVGMGNGTPCSLKQNRPRSSTVMYICHPESKHEILSVAEVTTC EYEVVILTPLLCSHPKYRFR  
ASPVNDIFCQSLPGSPFKPLTLRQLEQQEEILRVPFRNKEGVGWWKYEFCYGKHVHQYHEDK  
DSGKTSVVVG TWNQEEHIEWAKKNTARAYHLQDDGTQTVRMVSHFYGN GDICDITDKPRQVTV  
KLKCKESDSPHAVTVYMLEPHSCQYILGVESPVICKILD TADENGLLSLPN

**Important features of the protein:****Signal peptide:**

amino acids 1-30

**Glycosaminoglycan attachment site.**

amino acids 28-32

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 337-341

**N-myristoylation sites.**

amino acids 6-12, 23-29, 29-35, 49-55, 141-147, 152-158, 192-198,  
196-202

**Gram-positive cocci surface proteins 'anchoring' hexapeptide.**

amino acids 54-60

**FIGURE 95**

TTCCGTTTTCTGGGAGGAGTGAGGGGGCAACGGGTTCGGAGAAAAAAGGAAAAAAGAAGGGCTCAGC  
GCCTCCCCGCCGGGCCGTGGACAGAGGGGCACAGTTTTCGGCAGGCGGGTGAGGTTCGCTGAGGG  
CCCCCGGGAG**ATG**TTTTCTTGTCTGAGCACGGTGCAACCCAGGTTACAGTTCCTCTGAGTCA  
TCTCATCAATGCCTTCCATACACCAAAAAACACTTCTGTTTCTCTCAGTGGAGTGTCAAGTTTCT  
TCAAAACCAGCATCGAGATGTAGTTCCTGAGCATGAGGCTCCCAGCAGTGAGCCTTCACTTAA  
CTTAAGGGACCTTGGATTATCTGAACTAAAAATTGGACAGATTGATCAGCTGGTAGAAAAATCT  
ACTTCTCTGGATTTTGTAAAGGCAAAAAACATTTCTTCCCATTGGCATAACATCCCATGTCTCTG  
ACAATCCTTCTTTGAAAATAAATATGGTAACTTAGATATATTTAGTACATTACGTTTCTTCTG  
CTTGTATCGACATCATTCAAGAGCTCTTCAAAGCATTTGTTTCAGATCTTCAGTACTGGCCAGT  
TTTCATACAGTCTCGGGGTTTTAAACCTTTGAAATCAAGGACACGACGTCTCCAGTCTACCTC  
CGAGAGATTAGCTGAAACACAGAATATAGCGCCATCATTTCGTGAAGGGGTTTTCTTTTTCGGGA  
CAGAGGATCAGATGTTGAGAGTTTGGACAAACTCATGAAAACCAAAAAATATACCTGAAGCTCA  
CCAAGATGCATTTAAACCTGGTTTTTTCGGGAAGGTTTTCTGAAAGCTCAAGCACTCACACAAAA  
AACCAATGATTCCTTAAGGCGAACCCGTCTGATTCTCTTCGTTCTGCTGCTATTTCGGCATTTA  
TGGACTTCTAAAAAACCCATTTTTTATCTGTCCGCTTCCGGACAACAACAGGGCTTGATTCTGC  
AGTAGATCCTGTCCAGATGAAAAATGTCACCTTTGAACATGTTAAAGGGGTGGAGGAAGCTAA  
ACAAGAATTACAGGAAGTTGTTGAATTCTTGA AAAATCCACAAAAATTTACTATTCTTGGAGG  
TAAACTTCCAAAAGGAATTCTTTTAGTTGGACCCCCAGGGACTGGAAGACACTTCTTGCCCCG  
AGCTGTGGCGGGAGAAGCTGATGTTCTTTTTATTATGCTTCTGGATCCGAATTTGATGAGAT  
GTTTGTGGGTGTGGGAGCCAGCCGTATCAGAAATCTTTTTTAGGGAAGCAAAGGCGAATGCTCC  
TTGTGTTATATTTATTGATGAATTAGATTCTGTTGGTGGGAAGAGAATTGAATCTCCAATGCA  
TCCATATTTCAAGGCAGACCATAAATCAACTTCTTGCTGAAATGGATGGTTTTTAAACCCAATGA  
AGGAGTTATCATAATAGGAGCCACAAACTTCCCAGAGGCATTAGATAATGCCTTAATACGTCC  
TGGTCGTTTTTGACATGCAAGTTACAGTTCCAAAGGCCAGATGTAAAAGGTGCAACAGAAATTTT  
GAAATGGTATCTCAATAAAATAAAGTTTGATCAATCCGTTGATCCAGAAATTATAGCTCGAGG  
TACTGTTGGCTTTTTCCGGAGCAGAGTTGGAGAATCTTGTGAACCAGGCTGCATTTAAAGCAGC  
TGTTGATGGAAAAGAAATGGTTACCATGAAGGAGCTGGAGTTTTCCAAAGACAAAATTTCTAAT  
GGGGCCTGAAAGAAGAAGTGTGGAAATTGATAACAAAAACAAAACCATCACAGCATATCATGA  
ATCTGGTTCATGCCATTATTGCATATTACACAAAAGATGCAATGCCTATCAACAAAGCTACAAT  
CATGCCACGGGGGCCAACACTTGGACATGTGTCCCTGTTACCTGAGAATGACAGATGGAATGA  
AACTAGAGCCCAGCTGCTTGCACAAATGGATGTTAGTATGGGAGGAAGAGTGGCAGAGGAGCT  
TATATTTGGAACCGACCATATTACAACAGGTGCTTCCAGTGATTTTTGATAATGCCACTAAAAT  
AGCAAAGCGGATGGTTACCAAATTTGGAATGAGTGAAAAGCTTGGAGTTATGACCTACAGTGA  
TACAGGGAAACTAAGTCCAGAAACCCAATCTGCCATCGAACAAGAAATAAGAATCCTTCTAAG  
GGACTCATATGAACGAGCAAAAACATATCTTGAAAACCTCATGCAAAGGAGCATAAGAATCTCGC  
AGAAGCTTTATTGACCTATGAGACTTTGGATGCCAAAGAGATTCAAATTTGTTCTTGAGGGGGAA  
AAAGTTGGAAGTGAGAT**GA**TAACTCTCTTTGATATGGATGCTTGCTGGTTTTATTGCAAGAATA  
TAAGTAGCATTGCAGTAGTCTACTTTTTACAACGCTTTCCCCTCATTCTTGATGTGGTGTAATT  
GAAGGGTGTGAAATGCTTTGTCAATCATTTGTACATTTATCCAGTTTGGGTTATTCTCATT  
TGACACCTATTGCAAATTAGCATCCCATGGCAAATATATTTTGAAAAAATAAAGAAGTATCAG  
GATTGAAAACAAAAA

**FIGURE 96**

MFSLSSSTVQPQVTVPLSHLINAFTPKNTSVSLSGVSVSQNQHRDVVPEHEAPSSEPSLNLRD  
 LGLSELKIGQIDQLVENLLPGFCKGKNISSHWHTSHVSAQSFFENKYGNLDIFSTLRSSCLYR  
 HHSRALQSICSDLQYWPVFIQSRGFKTLKSRTRRLQSTSERLAETQNIAPSFVKGFLLRDRGS  
 DVESLDKLMKTKNIPEAHQDAFKTGFAEGFLKAQALTQKTNDSLRRTRLILFVLLLFGIYGLL  
 KNPFLSVRFRTTTGLDSAVDPVQMKNVTFEHVKGVEEAKQELQEVVEFLKNPQKFTILGGKLP  
 KGILLVGPPGTGKTLLARAVAGEADVPFYYASGSEFDEMFGVGASRIRNLFREAKANAPCVI  
 FIDELDSVGGKRIESPMHPYSRQTINQLLAEMDGFKPNEGVIIGATNFPEALDNALIRPGRF  
 DMQVTVPRPDVKGRTTEILKWYLNKIKFDQSVDPETIARGTVGFSGAELNVLNQAALKAAVDG  
 KEMVTMKELEFSKDKILMGPERRSVEIDNKNKTITAYHESGHAI IAYYTKDAMPINKATIMPR  
 GPTLGHVSLLPENDRWNETRAQLLAQMDVSMGGRVAEELIFGTDHITTGASSDFDNATKIAKR  
 MVTKFGMSEKLGVMYSDTGKLSPETQSAIEQEIRILLRDSYERAKHILKTHAKEHKNLAEAL  
 LTYETLDAKEIQIVLEGKKLEVR

**Important features of the protein:****Transmembrane domain:**

amino acids 238-259

**N-glycosylation sites.**

amino acids 28-32, 90-94, 230-234, 278-282, 535-539, 584-588,  
 623-627

**N-myristoylation sites.**

amino acids 35-41, 266-272, 286-292, 325-331, 357-363, 599-605

**Amidation site.**

amino acids 387-393, 709-713

**ATP/GTP-binding site motif A (P-loop).**

amino acids 322-330

**AAA-protein family proteins**

amino acids 315-336, 343-386, 405-451

**FIGURE 97**

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTTCTCCCCAGTTCCCCTGTGGGTCTGAGGG  
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATCCCCAA  
ACAAGTTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT  
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG  
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG  
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGT  
CTTGCCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG  
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT  
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACT  
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTGCTAAGACTC  
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC  
AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACA  
TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG  
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAG  
GAGACAGAAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTTCGAGGTCAAGAGCTCCAGTCT  
TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT  
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCAATGATTGTCTTTATGCATCCCC  
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT  
TATTAGTTAATATATTTATTTATTTTTTGTCTATTTAATGTATTTATTTTTTTTACTTGGACATG  
AACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTTAT  
ACAGTAAAAAAAAAAAAACCTTGTAATTCTAGAAGAGTGGCTAGGGGGGTTATTTCATTTGTAT  
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC  
TACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG  
ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG  
CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 98**

MKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNGFSEIRGSVQAKDGNIDIR  
ILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKISSLANSLTIKKDLRLC  
HAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEETE

**Signal sequence:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 107-110, 140-143

**N-myristoylation site.**

amino acids 51-56

**Interleukin 10:**

amino acids 9-176

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 99**

GCGCCGGCTCCGCGCCTCGCGCCCAGTCCGCGGGCCGCGCCGCCGCTCCCGCCGCTCCCGCCG  
 CTCCCGCAGCCGCCCCGCGCCCGCCCGGAGCCCCGCGTCCCTAGGCCTGGCTCCCGCCTGCC  
 CGAGACCCGCCCAGCCTGCCCCGCTCAGCCGCCAGAGAAG**ATG**CGGCTGCTCCCGGAATGGTT  
 CCTCTTGCTCTTTGGCCCGTGGCTCCTTAGGAAGGCCGTCAGTGCCAGATACCAGAGTCCGG  
 AAGGCCGCAGTACCTGGGGCTGCGCCCCGCGCGGCCGGAGCGGGTGCCCCCGGCCAGCAGCT  
 CCCAGAGCCAAGGTCTTCGGACGGCCTAGGCGTGGGCCGCGCCTGGAGCTGGGCCTGGCCGAC  
 CAACCACACGGGGGCGCTGGCCCCGGGCAGGGGCAGCCGGGGCGTTGCCCGCGCAGCGCACCAA  
 GAGGAAGCCGTCCATCAAGGCGGCGCGCGCCAAAAAGATCTTCGGCTGGGGGGACTTCTACTT  
 TCGGGTGCATACCCTCAAGTTTTTCGCTGCTGGTGACCGGCAAGATCGTGGACCATGTGAACGG  
 TACCTTCAGTGTGTATTTCCGCCACAACCTCGTCCAGCCTGGGCAACCTCAGTGTGAGCATCGT  
 GCCGCCCTCCAAGCGTGTGAGTTTCGGAGGAGTCTGGCTGCCCGGGCCTGTCCCCCACCCTCT  
 GCAGTCTACGCTCGCCCTGGAGGGGGTGCTTCCTGGGCTGGGGCCCCCGCTGGGGATGGCAGC  
 AGCAGCGGGCGGGGCGGGGCTTGGGGGCTCCCTCGGGGGCGCACTGGCGGGGGCCGCTTGGGGG  
 CGCGTTGGGAGTGCCCTGGGGCCAAAGAGTCACGCGCTTTCAATTGCCACGTGGAGTATGAGAA  
 GACAAACCGCGCGCGCAAGCACCGACCGTGCCTGTACGACCCGTCGCAGGTGTGTTTTACCGA  
 GCACACGCAGAGCCAGGCCGCCTGGCTCTGTGCCAAGCCCTTCAAAGTCATCTGTATCTTCGT  
 CTCTTTCCTCAGCTTTGACTACAACTGGTGCAGAAGGTGTGCCCAGACTATAACTTCCAGAG  
 TGAGCACCCCTACTTCGGA**TAG**CGCCCCCTCCCCAGCCAGTCCTGAGCCTCCCGCCAAATCCCA  
 GCCTCACTAGGTGGGACCCCCCTTCCAGTGTTCTGCCGCTCCTGTGGCCATGTCGCCCACTCC  
 TTCCACTCTGGGGGCGGAGGGGAATGGCTTCTCGGGACCCTCAGCTAGCGTGGGTGCCCTTTT  
 CCTTATGCGGAGTGCCCGCAAGGCTGGGGTAGCCCCCTCCAGTACACCCCAAAGTGAAAGGGA  
 TAAGAGTGCAGCCCCAGAATAGGCGGGGCTTGGAGGCGGTCCCAATGTCCCCTGGGTCCACAG  
 TGGGTCCCCTTTTCACCCTTGGCGCTAGGCTGCGCACTCCCTTTCCCCGCAGCTTTAATAACT  
 CCTGGCCTGGCACCCCTCACCCACCCCTGACTTTCCCATCCCCAGCGCTTGTCTTGCTTCACC  
 ATACCCCGCCTAAGACTGTAAAGGCCTAAAAACCTCGGCCTGTCCTCCCACCATTCTGCCTGC  
 CATATGCCTGTCCCCTTTTCCTCCAAACCCTATTAGGGTACCGGAAGCAGAACCCTGGGCTG  
 AGGCCCTGGCCCTGCCCCCGGCCCTGCCCTGCCCGCCCCCCTCCAGTCCAGGCAGTCGAGC  
 TCCACCTGCCCTCTCCTGCTGCTTCCTCTCGGTGATATTTTTTCTACGCCAAAACAGACGGGA  
 AAGGGAACAAAATAAAGTGAAATCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AA  
 AAAAAAAAAAAAAA

## **FIGURE 100**

MRLLPWFLLLFGPWLLRKAVSAQIPESGRPQYLGLRPAAAGAGAPGQQLPEPRSSDGLGVGR  
AWSWAWPTNHTGALARAGAAGALPAQRTKRKPSIKAARAKKIFGWGDFYFRVHTLKFSLLVTG  
KIVDHVNGTFSVYFRHNSSSLGNLSVSIVPPSKRVEFGGVWLPGPVPHPLQSTLALEGVLPGL  
GPPLGMAAAAAGPGLGGSLGGALAGPLGGALGVPGAKESRAFNCHVEYEKTNRARKHRPCLYD  
PSQVCFTEHTQSQAAWLCAKPFKVICIFVSFLSFDYKLVQKVC PDYNFQSEHPYFG

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **Transmembrane domain:**

amino acids 273-288

#### **N-glycosylation sites.**

amino acids 72-76, 133-137, 143-147, 149-153

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 93-97

#### **N-myristoylation sites.**

amino acids 35-41, 58-64, 60-66, 81-87, 84-90, 184-190, 194-200,  
203-209, 205-211, 206-212, 209-215, 217-223, 221-227, 224-230

#### **Cytochrome b/b6 Qo site signature.**

amino acids 5-11



FIGURE 101

AATGCCCC**ATG**CGCACCCACAGCTCGCGCTCCTGCAAGTGTTCTTTCTGGTGTTCCCCGATG  
GCGTCCGGCCTCAGCCCTCTTCCCTCCCCATCAGGGGCAGTGCCACGTCTTTGGAGCTGCAGC  
GAGGGACGGATGGCGGAACCCCTCCAGTCCCCTTCAGAGGCGACTGCAACTCGCCCCGGCCGTGC  
CTGGACTCCCTACAGTGGTCCCTACTCTCGTGACTCCCTCGGCCCTGGGAATAGGACTGTGG  
ACCTCTTCCCAGTCTTACCGATCTGTGTCTGTGACTTGACTCCTGGAGCCTGCGATATAAATT  
GCTGCTGCGACAGGGACTGCTATCTTCTCCATCCGAGGACAGTTTTTCTCCTTCTGCCTTCCAG  
GCAGCGTAAGGTCTTCAAGCTGGGTTTGTGTAGACAACCTCTGTTATCTTCAGGAGTAATTCCC  
CGTTTCCCTTCAAGAGTTTTTCATGGATTCTAATGGAATCAGGCAGTTTTTGTGTCCATGTGAACA  
ACTCAAACCTTAAACTATTTCCAGAAGCTTCAAAGGTCAATGCAACCAACTTCCAGGCCCTGG  
CTGCAGAGTTTTGGAGGCGAATCATTCACTTCAACATTCCAAACCTCAATCACCACCATCTTTTT  
ACAGGGCTGGGGACCCCATTTCTTACTTACTTCCCCAAGTGGTCTGTAATAAGCTTGCTGAGAC  
AACCTGCAGGAGTTGGAGCTGGGGGACTCTGTGCTGAAAGCAATCCTGCAGGTTTTCTTAGAGA  
GTAAAAGTACAACCTTGCACTCGTTTTTTTCAAGAACCTGGCTAGTAGCTGTACCTTGGAATTCAG  
CCCTCAATGCTGCCTCTTACTATAACTTCACAGTCTTAAAGGTTCCAAGAAGCATGACTGATC  
CACAGAATATGGAGTTCCAGGTTCTGTAATACTTACCTCACAGGCTAATGCTCCTCTGTTGG  
CTGGAAACACTTGTCAGAATGTAGTTTTCTCAGGTCACCTATGAGATAGAGACCAATGGGACTT  
TTGGAATCCAGAAAGTTTTCTGTCAAGTTTTGGGACAAACCAACCTGACTGTTGAGCCAGGCGCTT  
CCTTACAGCAACACTTCATCCTTCGCTTCAGGGCTTTTCAACAGAGCACAGCTGCTTCTCTCA  
CCAGTCCTAGAAGTGGGAATCCTGGCTATATAGTTGGGAAGCCACTCTTGGCTCTGACTGATG  
ATATAAGTTACTCAATGACCCTCTTACAGAGCCAGGGTAATGGAAGTTGCTCTGTTAAAAGAC  
ATGAAGTGCAGTTTGGAGTGAATGCAATATCTGGATGCAAGCTCAGGTTGAAGAAGGCAGACT  
GCAGCCACTTGCAAGCAGGAGATTTATCAGACTCTTCATGGAAGGCCAGACCAGAGTATGTTG  
CCATCTTTGGTAATGCTGACCCAGCCAGAAAGGAGGGTGGACCAGGATCCTCAACAGGCACT  
GCAGCATTTCACTATATAAAGTGTACTTCTGCTGTCTCATACCAAGTTTCCCTGGAGATCCAGG  
TATTGTGGGCATATGTAGGTCTCCTGTCCAACCCGCAAGCTCATGTATCAGGAGTTCGATTCC  
TATACCAAGTGCCAGTCTATACAGGATTCTCAGCAAGTTACAGAAGTATCTTTGACAACCTCTTG  
TGAACCTTTGTGGACATTACCCAGAAGCCACAGCCTCCAAGGGGCCAACCCAAAATGGACTGGA  
AATGGCCATTCGACTTCTTTCCCTTCAAAGTGGCATTCAGCAGAGGAGTATTCTCTCAAAAAT  
GCTCAGTCTCTCCCATCCTTATCCTGTGCCTCTTACTACTTGGAGTTCTCAACCTAGAGACTA  
TG**TGA**AGAAAAGAAAATAATCAGATTTCAAGTTTTCCCTATGAGAACTCTGAGGCAGCCACTT  
ATCTTGGCTAAATAGAACCTCACCTGCTCATGACCAGAGAGCATTTAGGATAATAGATGACCT  
AACTGAAGGAATCCTTGTATATGAAAGGAGTTATTTTAGAAAAGCAATAAAAATATTTTATTC  
ATCNTAAAAA

**FIGURE 102**

MRTPQLALLQVFFLVFPDGVRPQPSSSPSGAVPTSLELQRGTDGGTLQSPSEATATRPVAVPGL  
 PTVVPTLVTPSAPGNRTVDLFPVLPICVCDLTPGACDINCCCDRDCYLLHPRTVFSFCLPGSV  
 RSSSWVCVDNSVIFRSNSPFPSRVFMDNSGIRQFCVHVNNNSNLNYFQKLQKVNATNFQALAAE  
 FGGESFTSTFQTQSPPSFYRAGDPILTYFPKWSVISLLRQPAGVGAGGLCAESNPAGFLESKS  
 TTCTRFFKNLASSCTLDSALNAASYNFTVLKVPRSM TDPQNMEFQVPVILTSQANAPLLAGN  
 TCQNVVSQVTYEIETNGTFTGIQKVS VSLGQTNLTVEPGASLQQHFILRFRAFQQSTAASLTSP  
 RSGNPGYIVGKPLLALTDDISYSMTLLQSQNGSCSVKRHEVQFGVNAISGCKLRLKKADCSH  
 LQQEIYQTLHGRPRPEYVAIFGNADPAQKGGWTRILNRHCSISAINCTSCCLIPVSLEIQVLW  
 AYVGLLSNPQAHVSGVRFLYQCQSIQDSQQVTEVSLTTLVNFVDITQKPQPPRGQPKMDWKWP  
 FDFFPFKVAFSRGVFSQKCSVSPILILCLLLLGVNLNLETM

**Important features of the protein:****Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 484-505, 581-600

**N-glycosylation sites.**

amino acids 78-82, 165-169, 179-185, 279-285, 331-337, 347-351,  
 410-414, 487-491

**N-myristoylation sites.**

amino acids 30-36, 41-47, 124-130, 232-238, 236-242, 409-415

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 420-431



## **FIGURE 104**

MRRISLTSSPVRLLLFLLLLLLIALEIMVGGHSLCFNFTIKSLSRPGQPWCEAQVFLNKNLFLQ  
YNSDNNMVKPLGLLGKKVYATSTWGELTQTLGEVGRDLRMLLCDIKPQIKTSDPSTLQVEMFC  
QREAERCTGASWQFATNGEKSLLFDAMNMTWTVINHEASKIKETWKKDRGLEKYFRKLSKGDC  
DHWLREFLGHWPEAMPEPTGRRST

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 11-30 (possible type II protein)

#### **N-glycosylation site.**

amino acids 36-39, 154-157

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 2-5, 182-185, 209-212

#### **Casein kinase II phosphorylation site.**

amino acids 86-89, 93-96, 142-145, 185-188

#### **N-myristoylation site.**

amino acids 46-51

#### **Amidation site.**

amino acids 77-80, 207-210



**FIGURE 106**

MLAVSLTVPLL GAMMLLES PIDPQPLSFKEP LLLGVLHPNTKLRQAERLFENQLVGPESIAH  
IGDVMFTGTADGRVVKLENGEIETIARFGSGPCKTRDDEPVCGRPLGIRAGPNGTLFVADAYK  
GLFEVNPWKREVKLLLSSETPIEGKNMSFVNDLTVTQDGRKIYFTDSSSKWQRRDYLLLVMEG  
TDDGRLLLEYDTVTREVKVLLDQLRFPNGVQLSPAEDFVLVAETTMARIRRVYVSGLMKGGADL  
FVENMPGFPDNIRPSSSGGYWVG MSTIRPNPGFSMLDFLSERPWIKRMIFKLFSQETVMKFVP  
RYSLVLELSDSGAFRRSLHDPDGLVATYISEVHEHDGHL YLG SFRSPFLCRLSLQAV

**Important features of the protein:****Signal peptide:**

amino acids 1-13

**Transmembrane domain:**

amino acids 1-21 (possible type II)

**N-glycosylation sites.**

amino acids 116-119, 152-155

**Casein kinase II phosphorylation sites.**

amino acids 19-22, 27-30, 98-101, 146-149, 221-224, 286-289, 332-335

**N-myristoylation sites.**

amino acids 71-76, 92-97, 189-194, 244-249, 338-343

**Amidation site.**

amino acids 164-167

**FIGURE 107**

AACGAAGCGTGCGCGCTTTGGTAACCGGCTAGAAATCCCGCACGCGCGCCTGCCTCCTCTCCC  
CAGGCCTGAGCTGCCCTCCCACTGCCTTTCCTTCTTCCCGCGAGTCAGAAGCTTCGCGAGGG  
CCCAGAGAGGCGGTGGGGTGGGCGACCCTACGCCAGCTCCGGGCGGGAGAAAGCCCACCTCT  
CCCGCGCCCCAGGAAACCGCCGGCGTTTCGGCGCTGCGCAGAGCC**ATG**GAATTCTCCTGGCTGG  
AGACGCGCTGGGCGCGGCCCTTTTACCTGGCGTTTCGTGTTCTGCCTGGCCCTGGGGCTGCTGC  
AGGCCATTAAGCTGTACCTGCGGAGGCAGCGGCTGCTGCGGGACCTGCGCCCCCTTCCAGCGC  
CCCCACCCACTGGTTCCTTGGGCACCAGAAAGTTTATTTCAGGATGATAACATGGAGAAGCTTG  
AGGAAATTATTGAAAAATACCTCGTGCCTTCCCTTTCTGGATTGGGCCCTTTCAGGCATTTT  
TCTGTATCTATGACCCAGACTATGCAAAGACACTTCTGAGCAGAACAGATCCCAAGTCCCAGT  
ACCTGCAGAAATTCTCACCTCCACTTCTTGGAAAAGGACTAGCGGCTCTAGACGGACCCAAGT  
GGTTCAGCATCGTCGCCTACTAACTCCTGGATTCCATTTTAACATCCTGAAAGCATAACATTG  
AGGTGATGGCTCATTCTGTGAAAATGATGCTGGATAAGTGGGAGAAGATTTGCAGCACTCAGG  
ACACAAGCGTGGAGGTCTATGAGCACATCAACTCGATGTCTCTGGATATAATCATGAAATGCG  
CTTTCAGCAAGGAGACCAACTGCCAGACAAACAGCACCCATGATCCTTATGCAAAGCCATAT  
TTGAACTCAGCAAAATCATATTTACACGCTTGTACAGTTTGTGTATCACAGTGACATAATTT  
TCAAACCTCAGCCCTCAGGGCTACCGCTTCCAGAAGTTAAGCCGAGTGTGAAATCAGTACACAG  
ATACAATAATCCAGGAAAGAAAGAAATCCCTCCAGGCTGGGGTAAAGCAGGATAACACTCCGA  
AGAGGAAGTACCAGGATTTTCTGGATATTGTCTTTTCTGCCAAGGATGAAAGTGGTAGCAGCT  
TCTCAGATATTGATGTACACTCTGAAGTGAGCACATTCCTGTTGGCAGGACATGACACCTTGG  
CAGCAAGCATCTCCTGGATCCTTTACTGCCTGGCTCTGAACCCTGAGCATCAAGAGAGATGCC  
GGGAGGAGGTGAGGGGCATCCTGGGGGATGGGTCTTCTATCACTTGGGACCAGCTGGGTGAGA  
TGTCGTACACCACAATGTGCATCAAGGAGACGTGCCGATTGATTCTCCTGCAGTCCCGTCCATTT  
CCAGAGATCTCAGCAAGCCACTTACCTTCCCAGATGGATGCACATTGCCTGCAGGGATCACCG  
TGGTTCTTAGTATTTGGGGTCTTCACCACAACCCTGCTGTCTGGAAAAACCCAAAGGTCTTTG  
ACCCCTTGAGGTTCTCTCAGGAGAATTCTGATCAGAGACACCCCTATGCCTACTTACCATTCT  
CAGCTGGATCAAGGAAGTGCATTGGGCAGGAGTTTGCCATGATTGAGTTAAAGGTAACCATTG  
CCTTGATTCTGCTCCACTTCAGAGTGACTCCAGACCCACAGGCCTCTTACTTTCCCCAACC  
ATTTTATCCTCAAGCCCAAGAATGGGATGTATTTGCACCTGAAGAACTCTCTGAATGT**TAGA**  
TCTCAGGGTACAATGATTAAACGTACTTTGTTTTTCGAAGTTAAATTTACAGCTAATGATCCA  
AGCAGATAGAAAGGGATCAATGTATGGTGGGAGGATTGGAGGTTGGTGGGATAGGGGTCTCTG  
TGAAGAGATCCAAAATCATTTCTAGGTACACAGTGTGTCTAGCTAGATCTGTTTCTATATAACT  
TTGGGAGATTTTTCAGATCTTTTCTGTAAACTTTCACTACTATTAATGCTGTATACACCAATA  
GACTTTCATATATTTTCTGTTGTTTTTAAAATAGTTTTTCAGAATTATGCAAGTAATAAGTGCA  
TGTATGCTCACTGTCAAAAATTTCCCAACACTAGAAAATCATGTAGAATAAAAATTTTAAATCT  
CACTTCACTTAGCCGACATTCATGCCCTGACCAATCCTACTGCTTTTCTTAAACAGAATA  
ATTTGGTGTGCATTCTTTCAGACTTTTTTCTATACATTTTATATGTAGAAATGTAGCAATGTA  
TTTGTATAGATGTGATCATTCCTATATTGTTATTGATTTTTTTTCACTTAATAAAAATTCACCT  
TATTCCTTAAA

**FIGURE 108**

MEFSWLETRWARPFYLAFFVCLALGLLQAIKLYLRRQRLRLDLRPFPPAPPTHWFLGHQKFIQD  
 DNMEKLEEIIEKYPRAFPFWIGPFQAFFCIYDPDYAKTLLSRTDPKSQYLQKFSPPLLKGGLA  
 ALDGPKWQHRLLTPGFHFNILKAYIEVMAHSVKMMLDKWEKICSTQDTSVEVYEHINSMSL  
 DIIMKCAFSKETNCQTNSTHDPYAKAIFELSKIIFHRLYSLLYHSDIIFKLSPQGYRFQKLSR  
 VLNQYTDTI IQERKKS LQAGVKQDNTPKRKYQDFLDIVLSAKDESGSSFSDIDVHSEVSTFLL  
 AGHDTLAASISWILYCLALNPEHQERCREEVRGILGDGSSITWDQLGEMSYTTMCIKETCRLI  
 PAVPSISRDL SKPLTFPDGCTLPAGITVVLSIWGLHHNPAVWKNPKVFDPLRFSQENS DQRHP  
 YAYLPFSAGSRNCIGQEFAMIELKVTIALILLHFRVTPDPTRPLTFPNHFILKPKNGMYLHLK  
 KLSEC

**Important features of the protein:****Signal peptide:**

amino acids 1-29

**Transmembrane domains:**

amino acids 310-330, 397-413, 459-473

**N-glycosylation site.**

amino acids 206-210

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 265-269, 504-520

**N-myristoylation sites.**

amino acids 25-31, 298-304, 353-359, 450-456, 456-462

**Cytochrome P450 cysteine heme-iron ligand signature.**

amino acids 447-457

**Cytochrome P450 cysteine heme-iron ligand proteins.**

amino acids 444-475



**FIGURE 109**

GGCGTTCCGGGCCTCAACTTTGGCGTCGTGAGATTCTTGTGAGGCGTCTGCCTGGAAGCCGGC  
AGCAATTTTGCTTCTTTAAAGAGAAAAAGAAGGCTAGGGACTCAGATTCTTGATTCTGAGAT  
CCAGACCAGTCTCTCCAGACCTCTCCAGAAGAAGCC**ATG**GGAACCCCTCGTATCCAGCATTT  
GCTGATCCTCCTGGTCCTAGGAGCCTCCCTCCTGACCTCGGGCCTAGAGCTGTATTGTCAAAA  
GGGTCTGTCCATGACTGTGGAAGCAGATCCAGCCAATATGTTTAACTGGACCACAGAGGAAGT  
GGAGACTTGTGACAAAGGGGCACTTTGCCAGGAAACCATACTAATAATTAAAGCAGGGACTGA  
GACAGCCATTTTGGCCACGAAGGGCTGCATCCCGGAAGGGGAGGAGGCCATAACAATTGTCCA  
GCACTCTTCACCTCCCGGCCTGATCGTGACCTCCTACAGTAACTACTGTGAGGATTCCTTCTG  
TAATGACAAAGACAGCCTGTCTCAGTTTTGGGAGTTCAGTGAGACCACAGCTTCCACTGTGTC  
AACAAACCCTCCATTGTCCAACCTGTGTGGCTTTGGGGACCTGTTTCAGTGCTCCTTCTCTTCC  
CTGTCCCAATGGTACAACCTCGATGCTATCAAGGAAAACCTTGAGATCACTGGAGGTGGCATTGA  
GTCTGTGTGGAGGTCAAAGGCTGTACAGCCATGATTGGCTGCAGGCTGATGTCTGGAATCTT  
AGCAGTAGGACCCATGTTTGTGAGGGAAGCGTGCCACATCAGCTGCTCACTCAACCTCGAAA  
GACTGAAAATGGGGCCACCTGTCTTCCCATTCTGTTTGGGGGTTACAGCTACTGCTGCCATT  
GCTGCTGCCATCATTTATTCACTTTTCT**TAA**GAAGGCACCTTCTGGGCCTGGGTCTGAGGACAT  
CTTTTTTGAAGTGGGAGCCTTCTTACTGTTGAGGTTCAACAAGCTGAGGAGTAGATGGGAATTT  
GAGGGAGAATACAGAGATACTATGAACGTATTTGACATTTTAAATACAATTTCTGCTATAATT  
TTTGTATGCAGTAGGCGTTACTAATAAACATTTCTGCTGTGA

110/550

## **FIGURE 110**

MGTPRIQHLLILLVLGASLLTSGLELYCQKGLSMTVEADPANMFNWTTEEVEETCDKGALCQET  
ILIIKAGTETAILATKGCIPEGEEAITIVQHSSPPGLIVTSYSNYCEDSFENDKDSLSQFWEF  
SETTASTVSTTLHCPTCVALGTCFSAPSLPCPNGTTRCYQGKLEITGGGIESSVEVKGCTAMI  
GCRLMSGILAVGPMFVREACPHQLLTQPRKTENGATCLPIPVWGLQLLLPLLLPSFIHFS

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 184-201

#### **N-glycosylation sites.**

amino acids 45-49, 159-163

#### **N-myristoylation sites.**

amino acids 31-37, 70-76, 99-105, 147-153, 160-166, 174-180,  
175-181

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 111**

CCGAAGAGGACACAGAGGAGACTGAGCAAAGGGGGGTGGGCTCCAGGCGACCCCTAGCCCAATTCTGCCCTCCAT  
 CCCAAGGGGGCAGAGAAATTGTC'TTTCTTTGCTGACTCCTACGAGGAAAAAAAAAAAAAAAAAAACCATTATA  
 AGGGAAAGATAAACCGGAGACGGAGGAAAGGTGGCAGCCAGATTACTTAGAGAGGCACAGAGGAGAGAGATCGGGG  
 TGAGTCGCC**ATG**GGGACTCCCAGGGCCACGACCCGCCGCCTCCCAGCTGCTGTTCTTAATTCTGCTGAGCTGT  
 CCTTGGATCCAGGGTCTGCCCCCTGAAGGAGGAGGAGATATTGCCAGAGCCTGGAAGTGAGACCCCCACGGTGGCC  
 TCTGAGGCCCTGGCTGAAC'TGCTTACATGGGGCCCTGCTGAGGAGGGGCCAGAGATGGGTACCTGCCAGGATCT  
 GATCCGGACCCACGCTAGCCACCCCTCCGGCCGGCCAGACTCTCGCAGTGCCCTCCCTGCCACGGGCCACTGAG  
 CCGGGGACAGGGCCCTCTGACAACAGCCGTACCCCTAACGGGGTACGGGGGGCAGGCCCACTGCGCCAGAAGT  
 CTGACCCCGCCCCAGGAACCACAGCCCCACCCACCCAGCCCTGCCTCCCAGGGCCTCCCCTTGGGCCTGAG  
 GGAGGAGAGGAGGAGACGACGACCACCATCATCACCACGACAAC'TGTTACCACTACGGTGACCAGCCAGTTCTG  
 TGTATAACAACATCTCCGAGGGCGAAGGTATGTGGAGTCTCCAGATCTGGGGAGCCCCGTACGCCGACCCCTG  
 GGGCTCCTGGACTGCAC'TTAGCATCCATGTCTACCTGGCTACGGCATTGAGATCCAGGTGCAGACGCTGAAC  
 CTGTACAGGAAGAGGAGCTCCTGGTGCTGGCTGGTGGGGGATCCCCAGGCCTGGCCCCCGCACTCCTGGCCAAC  
 TCATCCATGCT'TGGAGAAGGACAAGTCTTTCGGAGCCCAACCAACCGGCTGCTTCTGCACTTCCAGAGCCCACGG  
 GTCCCAAGGGGCGGTGGCTTCAGGATCCACTATCAGGCCTACCTCCTGAGCTGTGGCTTCCCTCCCCGGCCGGCC  
 CATGGGGACGTGAGTGTGACGGACCTGCACCCCTGGGGGCACTGCCACCTTTCACTGTGATTTCGGGCTACCAGCTG  
 CAGGGAGAGGAGACCCCTCATCTGCC'TCAATGGCACCCCGGCCATCTTGAACGGTGAAACCCCCAGCTGCATGGCA  
 TCCTGTGTGGCACCATCCACAATGCCACCTGGGCCGCATCTGTCCCCAGAGCCTGGGGGAGCCGTAGGGCCC  
 AACCTACCTTCCGCTTGGGTTCATTGAAGCAGCTGAGGGGCGCCGGCTGCACCTGCAC'TTGAAGGGTCTCGCTG  
 GATGAGGACAATGACCGGCTGATGGTGCCTCAGGGGGCAGCCCCCTATCCCCCGTGATCTATGATTTCGGACATG  
 GACGATGTCCCCGAGCGGGGTCTCATAGTGACGCCCAGTCCCTCTACGTGGAGCTGCTGTGAGAGACACCTGCC  
 AATCCCCCTGCTGTTAAGCCTTCGATTTGAAGCCTTTGAGGAGGATCGCTGC'TTCGCCCCCTTCTGGCACATGGA  
 AATGTCAC'TACCACGGACCCCTGAGTATCGCCACGGGGCACTGGCAACCTTCTCGTGCCTCCCAGGATATGCCCTG  
 GAGCCCTTGGGCCCCCAATGCCATGTGTGGATCCCAAGAACCCCTGCGAACGACACAGAGCCGGCC  
 TGCAAAGCCATGTGTGGAGGGGAGCTGTGGAACAGAGCTGGCGTGGTCTCTCTCCCGACTGGCCCCAGAGCTAT  
 AGCCCCGGGCCAAGACTGCGTGTGGGGCGTGCACGTCCAGGAAGAGAAGCGCATCTTGCTCCAAGTTGAGATATTG  
 AATGTGCGGGAAGGGGACATGCTGACGCTGTTTCGACGGGGACGGTCCCAGCGCCCCGAGTCTTGGCCCAGCTGCGG  
 GGACCTCAGCCGCGCGCCGCTTCTCTCCTCTGGGCCCCGACCTCACACTGCAGTTTCAGGCACCGCCCGGGGCC  
 CCAATCCAGGCCCTGGGCCAGGGCTTCGTATTGCACTTCAAAGAGGTCCCGAGGAACGACAGCTGCCCCGAGCT  
 CCACTCCGGAGTGGGGCTGGAGAACCGCATCCACGGGGACCTGATCCGGGGCAGGTGCTCACCTACCACTGC  
 GAGCCTGGCTACGAGCTGTAGGCTCCGACATTTCTCACTTGCCAGTGGGACCTGTCTTGGAGCGCCGCGCCGCC  
 GCCTGCCAAAAGATCATGACTTGTGCTGACCTTGGCGAGATTGCCAACGGGACCCGACCCGCTCGGACGCCGGC  
 TTCCCCGTGGCTCCCACGTCCAGTACCGCTGCCTGCCAGGGTACAGCCTCGAGGGGGCAGCCATGCTCACCTGC  
 TACAGCCGGGACACAGGCACACCCAAGTGGAGCGATAGGGTCCCCAATGCGCCTTGAAGTACGAGCCGTGCCTG  
 AACCCGGGGGATCCCCGAGAATGGCTACCAAGCAGCTGTACAAGCACCATACCAAGCGGGGCGAGTCTCTGCGCTT  
 TCTGCTATGAGGGCTTGTAGCTTATCGCGAGGTACCATCACCTGTGTGCCCGGCCACCCCTCCAGTGGAC  
 AGCCAGCCCCCACTCTGCAAAGTGACCCAGACACAGATCCATCAGCGCAGCTGGAAGGGGGGAACCTGGCCCTG  
 GCCATCCTGCTGCCTCTAGGCTTGGTCATTGTCTCGGCAGTGGCGTTTACATCTACTACACCAAGCTTCAGGGA  
 AAGTCCCTTTTCGGCTTCTCGGGCTCCCACTCCTACAGCCCCATCACCGTGGAGTCGGACTTCAGCAACCCGCTG  
 TATGAAGCTGGGGATACGCGGGAGTATGAAGTTTCCATCT**TGA**ACCCCAAGACTACAGCTGCAGGACCCAGGACGC  
 CCTCCCTCTCATTCGGGCAGAGGGAAATACGGGACCCGGTCTCTGCCCTCCTGGCTGCCCTCCTCCCTGGCTG  
 TGTAAATAGTCTCCTTATCCACGAGGGGGCTTGTATGGCCCTGGAGATCCTACAGTAAATAAACCAGCATCCTG  
 CCGCCCAAAAAA

**FIGURE 112**

MGTPRAQHPPPPQLLFLILLSCPWIIQGLPLKEEEILPEPGSETPTVASEALAEELLHGALLRRG  
 PEMGYLPGSDDPDTLATPPAGQTLAVPSLPRATEPGTGPLTTAVTPNGVRGAGPTAPELLTTP  
 PGTTAPPPPPSPASPGPPLGPEGGEEETTTTIIITTTTVTSTVSPVLCNNNISEGEGYVESPD  
 GSPVSRTLGLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELLVLAGGSPGLAPRLLANS SMLG  
 EGQVLRSPNRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFPPRPAHGDVSVTDLHPGGTATFH  
 CDSGYQLQGEETLICLNGTRPSWNGETPSCMASC GGTIHNATLGRIVSPEPGGAVGPNLTCRW  
 VIEAAEGRRLLHLHFERSLDEDNDRMLMVRSGGSPLSPVIYDSMDDDVPERGLISDAQSLYVEL  
 LSETPANPLLLSLRFEAFEEEDRCFAPFLAHGNVT TTDPEYRPGALATFSCLPGYALEPPGPPN  
 AIECVDPTEPHWNDTEPACKAMCGGELSEPAGVVLSPDWPQSYSPGQDCVWGVHVQEEKRILL  
 QVEILNVREGDMLTLFDGDGPSARVLAQLRGPQPRRLLSSGPDLT LQFQAPPGPPNPGLGQG  
 FVLHFKEVPRNDTCPELPPPEWGWRTASHGDLIRGT VLT YQCEPGYELLGSDILT CQWDLSWS  
 AAPPACQKIMTCADPGEIANGHRTASDAGFPV GSHVQYRCLPGYSLEGAAMLTCYSRDTGTPK  
 WSDRVPKCALKEYEPCLNPGVPENGYQTL YKH HYQAGESLRFFCYEGFELIGEV TITCVP GHPS  
 QWTSQPPLCKVTQTTPDSRQLEGGNLALAILLPLGLVIVL GSGVYIYYTKLQGKSLFGFSGSH  
 SYSPITVESDFSNPLYEAGDTREYEVSI

**Important features of the protein:****Signal peptide:**

amino acids 1-27

**Transmembrane domain:**

amino acids 842-864

**N-glycosylation sites.**

amino acids 176-180, 222-226, 247-251, 332-336, 355-359, 373-377,  
 473-477, 517-521, 641-645

**Tyrosine kinase phosphorylation site.**

amino acids 61-69

**N-myristoylation sites.**

amino acids 2-8, 84-90, 111-117, 114-120, 190-196, 198-204,  
 235-241, 309-315, 333-339, 351-357, 472-478, 484-490, 528-534,  
 626-632, 665-671, 775-781, 842-848

**Amidation site.**

amino acids 384-388

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 12-23

**CUB domain proteins profile.**

amino acids 202-218, 376-392, 553-569

GCCGCGGGCGGAGCTGCTTGCCGGTCCCGCGCCGCGCGTCCGCACTCCTCGGCCCTCGGGCGGTGCGATGGGACGG  
 GCGCGCGCGGAGCAGGAGGCGCGCCCGTTCGGGGTGCTCGGGCCGCGCGGGAGCCCACTGTGGGGCTCGGGCATG  
 GCGGGCCGCGAGGACCTGAGCTCTCCTCAGGGGAGCGGGGAGGCAGCTGCTGGCCGGCGATGGGGACGGAGTGGGG  
 CCGTCGCCGCCGCGCCGAGCCGTGAGCGCCGAGCCACCGCCGCCGCTACCTCAGCCCTTCGCGAAGCGCCGGGCA  
 GCTCGGGAACATGCGCCCTGGAGCGGGCTCTGCTCGGTCTCTCAAAGTGTTGTTAATAACAGTACTGGTAGTGGAAGG  
 GATTGCGGTGGCCCCAAAAACCAAGATGGACAAAAATATTGGAATCAAGCATATTTCTGCAACCCAGTGTGGCAT  
 TTGGGTTGCAAGCAAAATGGAGGTCATTTTGTCTCGCAAATATTCCTGACTCATATCCACCAACAAGGAGTG  
 TATCTACATTTTGAAGCTGCTCCACGTCAAAGAATAGAGTTGACCTTTGATGAACATTATTATATAGAACCATC  
 ATTTGAGTGTGCGTTTGATCACTTGAAGTTCGAGATGGGCCATTTGGTTTTCTCTCTCTTATAGATCGTTACTG  
 TGGCGTGAAAAGCCCTCCATTAATTAGATCAACAGGGAGATTTCATGTGGATTAAAGTTTAGTTCTGATGAAGAGCT  
 TGAAGGACTGGGATTTTCGAGCAAAATATTCATTTATTCAGATCCAGACTTTACTTACCTAGGAGGTATTTTAAA  
 TCCCATTCCAGATTGTTCAGTTCGAGCTCTCGGGAGTCTGATGGAATAGTGCCTCTAGTCAGGTAGAACAAGAGGA  
 GAAAAACAAACAGGCCAAGCCGTTGATTGCATCTGGACATTAAAGCCATCCAAAGCTCAAGATTAAGATTTATTGAG  
 GTTCCTAGATTATCAAATGGAGCACTCAAATGAATGCAAGAGAAACTTCGTTGCAGTCTATGATGGAAGCAGTTTC  
 TATTGAAAACCTGAAGGCCAAGTTTTGCAGCACTGTGGCCAATGATGTAATGCTTAAACAGGAATTGGAGTGAT  
 TCGAATGTGGGCAGATGAAGGTAGTCGGCTTAGCAGGTTTTGAATGCTCTTTACTTCCCTTTGTGGAGCCTCCCTG  
 CACAAGCAGCACTTTCTTTTGCCATAGCAACATGTGCATCAATAATTCTTTAGTCTGTAATGGTGTCCAAAATTG  
 TGCATACCCCTTGGGATGAAAATCATTTGTAAAGAAAAAGAAAAAGCAGGAGTATTTGAACAAATCACTAAGACTCA  
 TGGAACAATTATTGGCATTACTTCCAGGATTTGCTTTGGTCTTTCTATTATTTCTATTTTTAGTACAAGTGAAACA  
 GCCTCGAAAAAAGGTCATGGCTTGCAAAAAACCGCTTTTATAAAACCGGGTTCCAAGAAGTGTTGATCCTCTCA  
 TTATGAACTGTTTTACTAAGGGACAAAGAGATTTCTGCAGACCTGGCAGACTTGTGCGAAGAAATTGGACAACTA  
 CCAGAAGATGCGGCGCTCCTCCACCGCCTCCCGCTGCATCCACGACCACCCTGTGGGTGCGAGGCCCTCCAGCGT  
 CAAACAAAGCAGGACCAACCTCAGTTCATGGAACCTCCTTTCCGAAATGACTTTGCACAACCACAGCCAATGAA  
 AACATTTAATAGCACCTTCAAGAAAAGTAGTTACACTTTCAAACAGGGACATGAGTGCCCTGAGCAGGCCCTGGA  
 AGACCGAGTAATGAGGAGATTTCCCTGTGAAATTTATGCTCAGGGGGCGAGAAGATTCTGCACAAGCATCCATATC  
 CATTGACTTCTAACTCTTCTGCTAATGGTGATGTGAATTTCTAGGGTGTGTACGTACGCGCCCTCCAGGCGCCATC  
 ACTGTTTCCAGCAGCCAACCCCTTTTCTCCCATCACAACTACGAAGACCTTGATTTACCGTTAACTTATTGTATGG  
 TGATGTTTTTATTCTCTCAGGCAGTCTATATATGTTAAACCAATCAAGGAACTTACTCTATTTCAGTGGAAACAAT  
 AATCATCTCTATTGCTTGGTGTCAATTTATAGGAAGCACTGCCAGTTAAAGAGCATTAGAAGAGGTGGTTGGATGG  
 AGCCAGGCTCAGGCTGCCTCTTTCGTTTTAGCAACAAGAAGACTGCTCTTGACTGATAACAGCTCTGTCAATATTT  
 TGATGCCACAATAAACTGATTTTTTTTTTACATTCCTTTTATTTTTCTTTCTCTAAATTTAATTTGTTTTATAA  
 GCCTATCGTTTTTACCATTTCATTTTCTTACATAAGTACAAGTGGTAAATGTACCACTACTTCAGTATAGGCATT  
 TGTTCTTGAGTGTGTCAAATACAGCTAGTTACTGTGCCAATTAAGACCCAGTTGTATTCTTACCCTATCTGTTTCT  
 TCTTGGCTAATCTCTGTACTTCTGCCTTTTAATTACTGGGCCCTTATTCCTTATTTTCTGTGAGAAATAATAGAT  
 GATATGATTTATTACCTTTCAATTATATTTTTCTCAGTTATACTAGAAAATTTCATAATCCTGGGATATATGTAC  
 CATTGTGAGCTATGACTAAAAATTTGAAAAAGATAAAAAATTTCTAGCAAGCCTTTGAAGTTTACCAAGTATAGTC  
 ACATTCAGTGACAGCCCATTCATTCAGTAAAGAATCATTTTCATTCACCTTTGGGAGAGGCCCTATAATTACATTTA  
 TTTGCAATGTTTTCTCTTCGCTAGATTGTTACATAGTCCCATTCTGTTGGTTTTGCTTACAGCATATGGTAACCA  
 AGGTTAGATGCCAGTTAAAAATTCCTTAGAAATGGATGAGCCTTGAGATTGCTTCTTAACCTGGGACATGACATTT  
 TTCTAGCTCTTATCAAGAATAACAACCTTCCACTTTTTTTTTTAAACTGCACCTTTTGACTTTTTTTTATGGTATAAAAA  
 CAATAATTTATAAACATAAAAGCTCATTGTGTTTTTTAGACTTTTTGATATTATTTGATACTGTACAAACTTTATT  
 AAATCAAGATGAAAGACCTACAGGACAGATTCCTTTTCAGTGTTACATCAGTGGCTTTGTATGCAAAATATGCTGT  
 GTTGGACCTGGACGCTATAACTTTATTGTAAAGACCTTGGAAATGTGGACATAAGCTCTTTCTTTCTTTTGTAC  
 TGATATTAGTTTGTGATAAATTTTTCACTGTGTGATATTTATGCTCTAAATCACTACACAAATCCCATAATAAAA  
 TATACATTGTCTACCTGAAAAAAA

**FIGURE 114**

MALERLCSVLKVLITVLVVEGIAVAQKTQDGQNIGIKHIPATQCGIWVRTSNGGHFASPNYP  
 DSYPNKECIYILEAAPRQRIELTFDEHYIIEPSFECRFDHLEVRDGPFGFSPLIDRYCGVKS  
 PPLIRSTGRFMWIKFSSDEELEGFRKAYSFIPDPDFTYLGILNPIPDQCQFELSGADGIVR  
 SSQVEQEEKTKPGQAVDCIWTIKATPKAKIYLRFLDYQMEHSNECKRNFVAVYDGSSSIENLK  
 AKFCSTVANDVMLKTGIGVIRMWADEGSRLSRFRMLFTSFVEPPCTSSTFFCHSNMCINNSLV  
 CNGVQNCAYPWDENHCKEKKKAGVFEQITKTHGTIIGITSGIVLVLLIISILVQVKQPRKKVM  
 ACKTAFNKTGFQEVFDPPHYELFSLRDKEISADLADLSEELDNYQKMRRSSTASRCIHDHHC  
 SQASSVKQSRTNLSSMELPFRNDFAPQPMKTFNSTFKKSSYTFKQGHECPEQALEDRVMEEI  
 PCEIYVRGREDQAQASISIDF

**Important features of the protein:****Signal peptide:**

amino acids 1-22

**Transmembrane domain:**

amino acids 348-369

**N-glycosylation sites.**

amino acids 311-315, 385-389, 453-457, 475-479

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 426-430, 479-483

**N-myristoylation sites.**

amino acids 22-28, 32-38, 54-60, 186-192, 279-285, 318-324,  
 348-354, 352-358, 441-447

**FIGURE 115**

GGTCTCTGTCCTTGGCTGTGGCTCCTGCGCTCTGGCTGAGCC**ATG**TTCTCTCTCGCCCTC  
CTCACTGAGCTTGGGAAGACTGCAAGCCCACGAAGGTTCTGAAGGAATATTTCTGCATGTCACA  
GTTCCACGGAAGATTAAGTCAAATGACAGTGAAGTTTCAGAGAGGAAGATGATTTACATCATT  
ACAATTGATGGACAACCTTACACTCTACATCTCGGAAAAACAATCATTCTTACCCCAGAACTTT  
TTGGTTTATACATATAATGAAACTGGATCTTTGCATTCTGTGTCTCCATATTTTATGATGCAT  
TGCCATTACCAAGGATATGCTGCCGAATTTCCAAATTCATTTGTGACACTCAGTATATGTTCT  
GGTCTCAGGGGATTTCTCCAGTTTGAAAAATATCAGTTATGGAATTGAACCAGTAGAATCTTCA  
GCAAGATTTGAGCATATAATTTATCAAATGAAAAATAATGATCCAAATGTATCCATTTTAGCA  
GTAAATTACAGTCATATTTGGCAGAAAGACCAGCCCTACAAAGTTCCTTTAAACTCACAGATA  
AAAAATCTTTCAAAACTATTACCCCAATATCTGGAAATATACATTATAGTGAAAAAGCTTTG  
ATGTTTACCCAGTTCAAATTGACTGTTATACTGTCTTCCTTGAATTGTGGTCAAATGAAAC  
CAGATTTCCACCAGTGGGGATGCTGATGATATATTACAAAGATTTTTGGCATGGAAACGGGAC  
TATCTCATCCTACGGCCCCATGACATAGCATACTTACTTGTTTACAGGAAACATCCTAAATAT  
GTGGGAGCAACATTTCTTGGCACC GTATGCAATAAAAGCTATGATGCAGGTATTGCTATGTAT  
CCAGATGCAATAGGTTTGAGGGGATTTTCGGTTATTATAGCTCAACTGCTTGGCCTTAATGTA  
GGATTAACATATGATGACATCACTCAGTGTCTGTCTGAGAGCTACATGCATCATGAATCAT  
GAAGCAGTGAGTGCCAGTGGTAGAAAGATTTTTAGCAACTGCAGCATGCACGACTATAGATAT  
TTTGTTCCAAATTTGAGACTAAATGCCTTCAGAAGCTTTCAAATTTGCAACCATTACATCAA  
AATCAACCAGTGTGTGGTAATGGGATTTTGGAATCCAATGAAGAATGTGACTGTGGTAATAAA  
AATGAATGTCAATTTAAGAAGTGCTGTGATTATAACACATGTAAACTGAAGGGCTCAGTAAAA  
TGTGGTTCTGGACCATGTTGTACATCAAAGTGTGAGTTGTCAATAGCAGGCACTCCATGTAGA  
AAGAGTATTGATCCAGAGTGTGATTTTACAGAGTACTGCAATGGAACCTCTAGTAATTGTGTT  
CCTGACACTTATGCACTGAATGGCCGTTTGTGCAAGTTGGGAACTGCCTATTGCTATAACGGA  
CAATGTCAAACACTACTGATAACCAGTGTGCCAAGATATTTGGAAAAGGTGCTCAAGGTGCTCCA  
TTTGCCTGTTTTAAAGAAGTTAATTCTCTGCATGAAAGATCTGAAAACGTGGTTTTTAAAAAT  
TCACAACCATTACCTTGTGAACGGAAGGATGTTCTCTGTGGAAAATTAGCTTGTGTTTCAGCCA  
CATAAAAATGCTAATAAAAGTGACGCTCAATCTACAGTTTATTCATATATTCAAGACCATGTA  
TGTGTATCTATAGCCACTGGTTCCTCCATGAGATCAGATGGAACAGACAATGCCTATGTGGCT  
GATGGCACCATGTGTGGTCCAGAAATGTACTGTGTAAATAAAACCTGCAGAAAAGTTCAATTTA  
ATGGGATATAACTGTAATGCCACCACAAAATGCAAAGGGAAAGGGATATGTAATAATTTTGGT  
AATTGTCAATGCTTCCCTGGACATAGACCTCCAGATTGTAAATTCAGTTTGGTTCCCCAGGG  
GGTAGTATTGATGATGGAAATTTTCAGAAATCTGGTGACTTTTATACTGAAAAAGGCTACAAT  
ACACACTGGAACAACTGGTTTTATTCTGAGTTTCTGCATTTTTCTGCCGTTTTTCATAGTTTTTC  
ACCACTGTGATCTTTAAAGAAATGAAATAAGTAAATCATGTAACAGAGAGAATGCAGAGTAT  
AATCGTAATTCATCCGTTGTATCAGAAAGCGATGACGTGGGACAT**TAA**TATTGCACAGAACTT  
CCATAGCAAATAACCTAAAGGAACGAATGTGCTTTATTTATAACCTTACGTTATCCCCAATGC  
ATTGTAAATGTCAAACCTTTTGGAAAATAAAGCCTGCGTGCCCTCCC

**FIGURE 116**

MFLLLALLTELGRLQAHEGSEGI FLHVTVP RKIKSNDSEVSE RKMIYIITIDGQPYTLHLGKQ  
 SFLPQNFLVYTYNETGSLHSVSPYFMMHCHYQGYAAEFNSFVTL SICSGLRGFLQFENISYG  
 IEPVESSARFEHIIYQMKNNDPNVSILAVNYSHIWQKDQPYKVPLNSQIKNLSKLLPQYLEIY  
 IIVEKALMFTQFKLTVILSSLELWSNENQISTSGDADDILQRFLAWKRDYLLILRPHDIA YLLV  
 YRKHPKYVGATFPGTVCNKSYDAGIAMYPDAIGLEGFSVIIAQLLGLNVGLTYDDITQCFCLR  
 ATCIMNHEAVSASGRKIFSNCSMHDYRYFVSKFETKCLQKLSNLQPLHQNPVCGNGILESNE  
 ECD CGNKNECQFKKCCDYNTCKLKGSVKCGSGPCCTSKCELSIAGTPCRKSIDPECDFTEYCN  
 GTSSNCVPD TYALNGRLCKLGTAYCYNGQCQT TDNQCAKIFGKGAQGAPFACFKEVNSLHERS  
 ENCGFKNSQPLPCERKDVLCGKLACVQPHKNANKSDAQSTVYSYIQDHVCVSIATGSSMRSDG  
 TDNAYVADGTMCGPEMYCVNKT CRKVHLMGYNCNATTKCKGKGICNNFGNCQCFPGHRPPDCK  
 FQFGSPGGSID DGNFQKSGDFYTEKGYNTHWNNWFILSFCIFLPFFIVFTT VIFKRNEISKSC  
 NRENAEYNRNSSSVVSESDDVGH

**Important features of the protein:****Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 665-684

**N-glycosylation sites.**

amino acids 36-39, 76-79, 122-125, 149-152, 156-159, 177-180,  
 270-273, 335-338, 441-444, 537-540, 587-590, 601-604, 703-706

**Casein kinase II phosphorylation sites.**

amino acids 74-77, 208-211, 221-224, 304-307, 337-340, 346-349,  
 376-380, 415-418, 499-502, 639-642, 708-711

**Tyrosine kinase phosphorylation site.**

amino acids 243-249

**N-myristoylation sites.**

amino acids 53-58, 79-84, 266-271, 298-303, 372-377, 403-408,  
 408-413, 442-447, 462-467, 469-474, 488-493, 567-572, 610-615,  
 616-621, 634-639

**Amidation site.**

amino acids 328-331



**FIGURE 117**

CCCACGCGTCCGCGGACGCGTGGGGCTCAGTGGGCGTCGCGCGAAGGCTAAGGGAGTGTGGCG  
GGCGGCTCCGGGAGCCAACATGCCTCGGTATGCGCAGCTGGTCATGGGCCCCGCGGGCAGCGG  
GAAGAGCACCTACTGTGCCACCATGGTCCAGCACTGTGAAGCCCTCAACCGGTCTGTCCAAGT  
TGTAACCTGGATCCAGCAGCAGAACACTTCAACTACTCCGTGATGGCTGACATCCGGGAACT  
GATCGAGGTGGATGATGTAATGGAGGATGATTCTCTGCGATTCCGTCCCAACGGAGGATTGGT  
ATTTTGCATGGAGTACTTTGCCAATAATTTTGGCTGGGAGTGTCTTGGCCATGTAGA  
GGACGACTATATCCTTTTTGATTGTCCAGGTCAGATTGAGTTGTACACTCACCTGCCTGTGAT  
GAAACATCTGGTCCAGCAGCTCGAGCAGTGGGAGTTCCGAGTCTGTGGAGTTTTTCTTGTTGA  
TTCTCAGTTCATGGTGGAGTCATTCAAGTTTATTTCTGGCATCTTGGCAGCCCTGAGTGCCAT  
GATCTCTCTAGAAATTCCGCAAGTCAACATCATGACAAAAATGGATCTGCTGAGTAAAAAAGC  
AAAAAAGGAAATTGAGAAATTTTTAGATCCAGACATGTATTCTTTATTAGAAGATTCTACAAG  
TGACTTAAGAAGCAAAAAATTCAAGAACTGACTAAAGCTATATGTGGACTGATTGATGACTA  
CAGCATGGTTCGATTTTTACCTTACGATCAGTCAGATGAAGAAAGCATGAACATTGTATTGCA  
GCATATTGATTTTGCCATTCAATATGGAGAAGACCTAGAATTTAAAGAACCAAAGGAACGTGA  
AGATGAGTCTTCCTCTATGTTTGACGAATATTTTCAAGAATGCCAGGATGAATTGAAGAGTTTA  
CTAAAAGTAACCATCTAAAGAGCTTGTGGCCAAACCAGCAGAACATTCTTCTCTTCAAAGGAT  
GCAATAGTAGAAAGCTACTTATTTTAATGAAAAAAAGTAAACTTCGTTCTTTATCAGCCTCA  
TGCCTGAATCAAATTTTTAATTATTCTGAACTGCTGCTGTTTAAAGTGGAATCTTTTAGTAT  
TATAACAGCATCACTTTAGATTTTGTAAGTCAAAATTGAAATGAATGCACATAGATTTATATA  
TAAATTAGCACCTGAGCTAAGGTAAAGGCCGGTCTAACTTATTTTCACTTTTTGTATTATTT  
TTGAGATGCAGGAATTACTGTAACAAAATATGTATGTCCGAAGGGAAAAAGCTGCAAGGATAT  
ATATAAGACCACTGCTTATCTGTATCTTCCCATTTTCCTATATTGAAAATGTATATTATTTAT  
ATAACTTAAAAAGTAAAAATAACTATGTTTTGAGAT

**FIGURE 118**

MPRYAQLVMGPAGSGKSTYCATMVQHCEALNRSVQVVNLDPAAEHFNYSVMADIRELIEVDDV  
MEDDSLRFPGNGGLVFCMEYFANNFDWLENCLGHVEDDYILFDCPGQIELYTHLPVMKHLVQQ  
LEQWEFRVCGVFLVDSQFMVESFKFISGILAAALSAMISLEIPQVNIMTKMDLLSKKAKKEIEK  
FLDPDMYSLLEDSTSDLRSKFKKLTKAICGLIDDYSMVRFLPYDQSDEESMNIVLQHIDFAI  
QYGEDLEFKEPKEREREDESSMFDEYFQECQDE

**Important features of the protein:****Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 151-170

**N-glycosylation sites.**

amino acids 31-35, 47-51

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 212-216

**Tyrosine kinase phosphorylation site.**

amino acids 189-197

**N-myristoylation sites.**

amino acids 13-19, 76-82, 154-160

**ATP/GTP-binding site motif A (P-loop).**

amino acids 10-18

# FIGURE 119

GGGCGCTGGGAGACACCGGACGCGCCGCTCGGCTGCGCTGCGGCTCAGGCCCGCTCGGGCC  
GACCCGCTCGGTACCGCCGGCTCGGGCGCGCACCTGCCGGCTGCGCCCCAGGGCTCATGCGG  
AGGCCACGAGGAGGCCGGCGGCCACGCGCATCCCGTAGCCAGGTGGCCAGGTCTGCACCG  
CGGCGGCTCGGCGCC**ATG**GAGCCCCCGTATTTCGCTGACGGCGCACTACGATGAGTTCCAAGA  
GGTCAAGTACGTGAGCCGCTGCGGCGCGGGGGGCGCGCGCGGGGCTCCCTGCCCCGGGCTT  
CCCGTTGGGCGCTGCGCGCAGCGTCACCGGGGCCCGGTCCGGGCTGCCGCGCTGGAACCGGCG  
CGAGGTGTGCTGCTGTCGGGGCTGGTGTTCGCCGCCGGCTCTGCGCCATTCTGGCGGCTAT  
GCTGGCCCTCAAGTACCTGGGCCCCGGTTCGCGGCCGGCGGGCGGCGCTGTCCCGAGGGCTGCC  
TGAGCGCAAGGCCTTCGCGCGCGCCGCTCGCTTCCTGGCCGCCAACCTGGACGCCAGCATCGA  
CCCATGCCAGGACTTCTACTCGTTTCGCTGCGGCGGTTGGCTGCGGCGCCACGCCATCCCCGA  
CGACAAGCTCACCTATGGCACCATCGCGGCCATCGGCGAGCAAAACGAGGAGCGCCTACGGCG  
CCTGCTGGCGCGGCCCGGGGGTGGGCTGGCGGCGCGGCCACGCGCAAGGTGCGCGCCTTCTT  
CCGCTCGTGCTCGACATGCGCGAGATCGAGCGACTGGGCCCGGACCCATGCTAGAGGTGAT  
CGAGGACTGCGGGGCTGGGACCTGGGCGGCGCGGAGGAGCGTCCGGGGGTGCGGCGCGGATG  
GGACCTCAACCGGCTGCTGTACAAGGCGCAGGGCGTGTACAGCGCCGCGCGCTCTTCTCGCT  
CACGGTCAGCCTGGACGACAGGAACCTCTCGCGCTACGTATCCGCAATTGACCAGGATGGGCT  
CACCTGCCAGAGAGGACCTGTACCTCGCTCAGGATGAGGACAGTGAGAAGATCCTGGCAGC  
ATACAGGGTGTTTCATGGAGCGAGTGCTCAGCCTCCTGGGTGCAGACGCTGTGGAACAGAAGGC  
CCAAGAGATCCTGCAAGTGGAGCAGCAGCTGGCCAACATCACTGTGTGAGAGTATGACGACCT  
ACGGCGAGATGTCAGCTCCATGTACAACAAGGTGACGCTGGGGCAGCTGCAGAAGATCACCCC  
CCACTTGCGGTGGAAGTGGCTGCTAGACCAGATCTTCCAGGAGGACTTCTCAGAGGAAGAGGA  
GGTGGTGCTGCTGGCGACAGACTACATGCAGCAGGTGTGCGAGCTCATCCGCTCCACACCCCCA  
CCGGGTCTCTGCACAACTACCTGGTGTGGCGCGTGGTGGTGGTTCCTGAGTGAACACCTGTCCCC  
GCCATTCCGTGAGGCACTGCACGAGCTGGCACAGGAGATGGAGGGCAGCGACAAGCCACAGGA  
GCTGGCCCCGGTCTGCTTGGGCCAGGCCAATCGCCACTTTGGCATGGCGCTTGGCGCCCTCTT  
TGTACATGAGCACTTCTCAGCCGCCAGCAAAAGCCAAGGTGCAGCAGCTAGTGGAAGACATCAA  
GTACATCCTGGGCCAGCGCCTGGAGGAGCTGGACTGGATGGACGCCGAGACCAGGGCTGCTGC  
TCGGGCCAAGCTCCAGTACATGATGGTGATGGTTCGGCTACCCGGACTTCTGCTGAAACCCGA  
TGCTGTGGACAAGGAGTATGAGTTTGAGGTCCATGAGAAGACCTACTTCAAGAACATCTTGAA  
CAGCATCCCCCTCAGCATCCAGCTCTCAGTTAAGAAGATTCCGCAGGAGGTGGACAAGTCCAC  
GTGGCTGCTCCCCCACAGGCGCTCAATGCCTACTATCTACCCAACAAGAACCAGATGGTGT  
CCCCGCGGGCATCCTGCAGCCACCCCTGTACGACCCTGACTTCCCACAGTCTCTCAACTACGG  
GGGCATCGGCACCATCATTGGACATGAGCTGACCCACGGCTACGACGACTGGGGGGGCCAGTA  
TGACCGCTCAGGGAACCTGCTGCACTGGTGGACGGAGGCCTCCTACAGCCGCTTCTGCGAAA  
GGCTGAGTGCATCGTCCGTCTCTATGACAACTTCACTGTCTACAACCAGCGGGTGAACGGGAA  
ACACACGCTTGGGGAGAACATCGCAGATATGGGCGTCCTCAAGCTGGCCTACCACGCCTATCA  
GAAGTGGGTGCGGGAGCACGGCCAGAGCACCCACTTCCCCGGCTCAAGTACACACATGACCA  
GCTCTTCTTCATTGCCCTTTGCCCAGAACTGGTGCATCAAGCGGCGGTGCGAGTCCATCTACCT  
GCAGGTGCTGACTGACAAGCATGCCCTGAGCACTACAGGGTGCTGGGCAGTGTGTCCAGTT  
TGAGGAGTTTGGCCGGGCTTTCCACTGTCCCAAGGACTCACCCATGAACCCTGCCACAAGTG  
TTCCGTGTGG**TGA**GCCTGGCTGCCCCGCTGCACGCCCCCACTGCCCCCGCACGAATCACCTCC  
TGCTGGCTACCGGGGCAGGCATGCACCCGGTGCCAGCCCCGCTCTGGGCACCACCTGCCTTCC  
AGCCCCCTCAGGACCCGGTCCCCCTGCTGCCCCCTCACTTCCAGGAGGGGCTGGAGCAGGGTGA  
GGCTGGACTTTGGGGGGCTGTGAGGGAAATATACTGGGGTCCCCAGATTCTGCTCTAAGGGGG  
CCAGACCCTCTGCCAGGCTGGATTGTACGGGCCCCACCTTCGCTGTGTTCTTGCTGCAAAGTC  
TGGTCAATAAATCACTGCACTGTTAAAAA

**FIGURE 120**

MEPPYSLTAHYDEFQEVKYVSRGAGGARGASLPPGFPLGAARSVTGARSGLPRWNRREVCLL  
 SGLVFAAGLCAILAAMLALKYLGPVAAAGGGACPEGCPERKAFARAARFLAANLDASIDPCQDF  
 YSFACGGWLRRAIPDDKLTGTIAAIGEQNEERLRLLARPGGGPGGAAQRKVRAFFRSCLD  
 MREIERLGPRPMLIEVEDCGGWDLGGAERPGVAARWDLNRLLYKAQGVYSAAALFSLTVSLD  
 DRNSSRYVIRIDQDGLTLPERTLYLAQDEDSEKILAAAYRVFMERVLSELLGADAVEQKAQEILQ  
 VEQQLANITVSEYDDLRRDVSSMYNKVTLGQLQKITPHLRWKWLLDQIFQEDFSEEEEVVLLA  
 TDYMQQVSQILIRSTPHRVLHNYLVWRVVVVLSEHLSPPFREALHELAQEMEGSDKPPQELARVC  
 LGQANRHFGMALGALFVHEHFSAASKAKVQQLVEDIKYILGQRLEELDWMMDAETRAAARAKLQ  
 YMMVMVGYPDFLLKPDVADKEYEFEFVHEKTYFKNILNSIPFSIQLSVKKIRQEVDKSTWLLPP  
 QALNAYYLPNKNQMVFPAGILQPTLYDPDFPQSLNYGGIGTIIIGHELTHGYDDWGGQYDRSGN  
 LLHWWTEASYSRFLRKAECIVRLYDNFTVYNQRVNGKHTLGENIADMGVCLKLAYHAYQKWVRE  
 HGPEHPLPRLKYTHDQLFFIAFAQNWCICKRRSQSIYLQVLTDKHAPEHYRVLGSVVSQFEEFGR  
 AFHCPKDSPMNPAAHKCSVW

**Important features of the protein:****Transmembrane domain:**

amino acids 64-88

**N-glycosylation sites.**

amino acids 255-259, 322-326, 656-660

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 722-726

**N-myristoylation site.**

amino acids 24-30, 26-32, 27-33, 40-46, 47-53, 65-71, 148-154,  
 169-175, 170-176, 237-243, 450-456, 604-610, 607-613

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 85-96

**Prenyl group binding site.**

amino acids 772-777

**Neutral zinc metallopeptidases, zinc-binding region signature.**

amino acids 609-619

**Neutral zinc metallopeptidases, zinc-binding region proteins.**

amino acids 609-619

FIGURE 121

CGGACTGCCCGGACCGCGCGATGGAGTTCGACCGGCAGCGTCGGGGAGGCCCCCGGGCGGACCCC  
GGGTGCTGGTGGTGGGCGGCGGCATCGCGGGGCTGGGCGCGGCGCAGAGGCTCTGCGGCCACT  
CCGCCTTCCCGCACCTGCGGGTCCTGGAGGCCACGGCCCGCGCCGGGGGCCGCATCCGCTCGG  
AGCGCTGCTTCCGTGGCGTGGTGGAGGTGGGCGCGCACTGGATCCATGGGCCCTCCCGGGGTA  
ACCCCGTCTTCCAGCTGGCTGCTGAGTACGGGCTGCTGGGGGAGAAGGAGCTGTCCCAGGAGA  
ACCAGCTGGTGGAGACCGGGGGTACGTGGGCCTGCCCTCCGTGAGCTACGCCAGCTCCGGGG  
CCAGCGTGAGCCTCCAGCTGGTGGCGGAGATGGCGACTCTGTTCTACGGCCTGATAGACCAGA  
CCCGGGAGTTCCTGCACGCTGCAGAGACCCCGGTGCCAGCGTCGGGGAGTACCTCAAGAAGG  
AGATTGGCCAGCACGTGGCCGGCTGGACAGAGGATGAGGAGACCAGGAAGCTGAAGCTGGCCG  
TCCTGAACCTCCTTCTTCAACCTGGAATGCTGTGTGAGCGGCACCCACAGCATGGACCTGGTGG  
CCCTGGCACCCCTTTGGGGAGTATACCGTGCTGCCGGGGCTGGACTGCACCTTTTCTAAGGGCT  
ATCAAGGACTCACAACTGCATGATGGCCGCCCTGCCGGAGGACACTGTAGTTTTTTGAGAAGC  
CTGTGAAGACCATCCACTGGAACGGGTCCCTCCAGGAGGCAGCCTTTCCCGGGGAGACCTTTC  
CAGTGTCGGTAGAGTGTGAGGATGGAGACCGGTTCCCGGCGCACCATGTCATCGTCACCGTGC  
CCTTAGGTTTTTCTTAGGGAACATTTGGACACCTTCTTTGACCCTCCCCTGCCGGCTGAGAAGG  
CAGAAGCAATCAGGAAGATAGGCTTTGGGACCAACAACAAAATCTTCCTGGAGTTTGAGGAGC  
CCTTCTGGGAGCCAGACTGCCAGCTGATCCAGCTGGTGTGGGAGGACACGTCGCCCCCTGGAGG  
ATGCTGCCCCCTGAGCTACAGGACGCCTGGTTCGGGAAGCTCATTGGCTTTGTGGTCCTGCCTG  
CCTTTGCGTCTGTCCACGTTCTCTGTGGGTTCAATTGCCGGACTTGAGTCTGAGTTCATGGAGA  
CTCTGTCGGATGAAGAAGTACTTCTGTGTCTCACCCAAGTGCTCCGGAGAGTGACAGGAAACC  
CACGGCTCCCCGCGCCCAAGAGCGTCTGCGGTCTCGCTGGCACAGCGCCCCGTACACTAGGG  
GGTCCTACAGCTACGTGGCCGTGGGCAGTACTGGGGGCGACCTGGACCTGCTGGCTCAGCCCC  
TCCCTGCAGACGGCGCCGGCGCCAGCTCCAGATCCTGTTTGCGGGGGAAGCCACACATCGCA  
CGTTTTACTCCACGACGCACGGGGCTCTGCTGTCGGGATGGAGGGAGGCCGACCGCCTCCTCA  
GTCTGTGGGCCCCGCAGGTGCAGCAGCCCAGGCCGAGGCTCTAGCTGGGCCCAGCCTACTCTG  
TTCCACCCGTGTCGGGGGTAGGCTGGGACCGTCATTTCTTCTGACAGATTTTCACTGCTGGCTTG  
AAATTTGGGGATGTTAATGAGGGTCCTCTGGTTTTTTGGTAACCAGGGCCACCTTCTCAGTTCT  
TGTGTCTGTTATTGGAGTCTGGCCAGGGTTGACTTGAGCTGAGACACCAGATGCTCACGGAGA  
TGCTGGACACATAAAGCAAGTTACAGCCACAAAAAAAAAAAAA

**FIGURE 122**

MESTGSVGEAPGGPRVLVVGGGIAGLGAAQRLCGHSAFPHLRVLEATARAGGRIRSERCFGGV  
VEVGAHWIHGPSRGNPVFQLAAEYGLLGEKELSQENQLVETGGHVGLPSVSYASSGASVSLQL  
VAEMATLFYGLIDQTREFLHAAETPVPSVGEYLLKKEIGQHVAGWTEDEETRKLKLAVLNSFFN  
LECCVSGTHSMDLVALAPFGEYTVLPGLDCTFSKGYQGLTNCMAALPEDTVVFEKPVKTIHW  
NGSFQEAAFPGETFPVSVECEDGDRFPAHHVIVTVPLGFLREHLDTFFDPPLPAEKAEAIRKI  
GFGTNNKIFLEFEEPFWEPCQLIQLVWEDTSPLEDAAPELQDAWFRKLIGFVVLPAFASVHV  
LCGFIAGLESEFMETLSDEEVLLCLTQVLRRTGNPRLPAPKSVLRSRWHSAPYTRGSYSYVA  
VGSTGGDLDLLAQPLPADGAGAQLQILFAGEATHRTFYSTTHGALLSGWREADRLLSLWAPQV  
QQPRPRL

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 364-385

**N-glycosylation site.**

amino acids 253-257

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 408-412

**N-myristoylation sites.**

amino acids 20-26, 21-27, 25-31, 105-111, 119-125, 164-170,  
216-222, 227-233, 443-449, 484-490

**Aminooxidase Flavin containing amine oxidase:**

amino acids 23-497

# FIGURE 123

CGAGCGCGTGGGGGAAGAGGGATAAATAATTCTGTGCACACGTGCCCTGGCCTCTGGAGCTCAGCTGCCAGTCCAC  
GTCTAGGGAATCTTAGCATCTGGGACCAAGACACTTTACAGCAATCATCACCCCTTTGCAGAGGAGGTGAGCTCAC  
CAGGACTCATCTGCCATTTTCAGACCTTTTGTCTGTACCTGCCAGGTGGCCCCCAGCTGCTGACGAGAGATGATGGTGA  
TCTCTAGTCTCCCCGACTCCTTGAAGCCAGTATCGCTGACCGAGCTTGTCTTCTCATGACACCTCTCTCTC  
CCTTCAGCCTGGGGAGCCGAGCTCAGAGGTC AAGGTGCTAGGCCCTAGTATCCCATCCTGGCCCTCGTCGGGGA  
GGAGGTGGAGTTCCCCTGCCACCTATGGCCACAGCTGGATGCCAGCAAATGGAGATCCGCTGGTTCCGGAGTCA  
GACCTTCAATGTGGTACACCTGTACCAGGAGCAGCAGGAGCTCCCTGGCAGGCAGATGCCGGCGTTCCGGAACAG  
GACCAAGTTGGTCAAGGACGACATCGCCTATGGCAGCGTGGTCTGCAGCTTACAGCATCATCCCCCTCTGACAT  
GGGCACATATGGCTGCCGCTTCCACTCCGACAACCTTCTCTGGCGAAGCTCTCTGGGAACTGGGAGGTAGCAGGGC  
GGGCTCAGACCCCTACCTCTCCCTTGAAGGCTTCAAGGAAGGAGGCACTTACGTGAGGCTCAGATCAGTGGGCTG  
GTACCCCAAGCCTAAGGTTCAAGTGGAGAGACCACAGGGACAGTGCCTGCCCTCAGAGTTTGAAGCCATCGTCTG  
GGATGCCCAGGACCTGTTCACTGTGGAAACATCTGTGGTTGTCCGAGCGGGAGCCCTCAGCAATGTGTCCGTCTC  
CATCCAGAATCTCTCTTGAAGCCAGAAAGAAAGAGTTGGTGGTCCAGATAGCAGACGTGTTTCGTACCCGGAGCCTC  
TGCGTGGAAAGAGCGCTTCTGTCGCGACCCCTGCCGCTGCTGTTGGTCTTCGCGGCGCTGGCGCTGGGCGTCTCCG  
GAAGCAGCGGAGAAGCCGAGAAAGCTGAGGAAGCAGCGCGAGAAGAGACAAGAAACTCACTGCAGAGCTGGA  
AAAGCTTCAGACAGAGCTTGCATGGAGACGGGCTGAAGGCCAGGCTAGTGGAGAGCAGCCAAAAATATGCACT  
GGATGTGACGCTGGACCCGGCCTCGGCGCACCCAGCCTGGAGGTGTCCGAGGATGGCAAGAGCGTGTCTTCCCG  
CGGGGCGCCGCCAGGCCCGGCGCTGGCCACCCGCAGCGGTTCTCGGAGCAGACGTGCGCGCTGAGCCTGGAGCG  
GTTCTCCGCGCGCCGCTACTCTGGGAGGTGCACGTGGCGCGCCGCGAGCCGCTGGTTCTTGGGCGCCTGCCCTGGC  
CGCGTGGCGCGCGCGGGGCTGCGCGCTGAGCCCTGCGCGCGCTACTGGTGTCTGGGCTGTGGAACGGCTG  
CGAGTACTTCTCTGGCCCCGACCGCGTGCCTCACCTTGCCTGCGCGTCCCGCGCGCGCTGGGCGTCTCTCT  
GGACTACGAGGCCGGAGAGCTGTCTTCTTCAACGTGTCCGACGGCTCCACATCTTCACTTCCACGACACCTT  
CTCGGGCGCGCTCTGTGCGTACTTCAGGCCCAGGGCCACGACGGCGGCGAACATCCGGATCCCCTGACCATCTG  
CCCCGTGCGCGTTAGAGGGACGGCGCTCCCCGAAGAGAACGACAGTGCACCTGGCTACAGCCCCCTATGAGCCCCG  
GGACCCGCCCTGGACTGGTGGATGAGGCGCCCTCGTGGCGCGGAGCTGGCCCCGGGGGGCCCCCTGGATCCCCG  
GCCAGCGCTTGTCTCTCTGCTCGCTGTAAGGGAGCGAGGTGCACCGCAAATGTGACGAGGGGGGACAAAGA  
GAGGGACCTTTGCCCTACGTAGATGTGTATGTGTAGTGCATTTTCTTCAAGGAAAGGAGACAAGTCCAAAGCTCG  
TTTGTGGATTGTGGGACTGAGCGAAGGAGTACAAATATATCCACGTCGCTCAGAGCTGGGGTGCTCACGGTGGGC  
GGTGGGCAAGAAAGCCAGCATGGAAGAAAGAAAGGGAGAAAACCTTTGGTGACTGCCTTAGAGGGATCAGTTAATTTG  
TATAGTTTTATATTTTTGTATATGTTTGTAGCTTAAAGGTCAGATGCAATAACACTTCGTAAGCAACGA  
GTTCACTTAAGTAAGCTCAGATCCTAGTTTAAAAAACCTTTCCCTATAAATGAAGTTGGAGGAACAGCTGCT  
TCTGAGCCGGGGCAAAAAATTTCAAGGTGAGCCTGGAGCATTTGTGTGGTGAAGTAAAAATAAAGGCTCAAACGT  
GACGGCAACCCGGCAAAGGGTAGGGAGCCAGGCCGAAGGGGCTCACTGACCAATTGTGGGACAATTTGAACAT  
CAGGATGAATAATGACAGGAGAGATTATAACACACTGAATAAAACATAATCCATGAGTTTCATGCTGATACCTCAA  
ATTTCTTTTTTAAAAAGGAGAAACAGGAAGTTCTTTTGGAGGTGAAATCTAATTATTTGGTGAGAGTCTTTGGGA  
ACAGCGTGTTCAGCTCTCAAAGCAGTAACCTTATACACTACTTATAAGTTTGAAGGGGAAAGGTTACCTTTTAC  
AATGGAGACATCTACCAGATCATCCAAGTGATTAAATTTAACATCATCAATGATGGGACCAAGGACATTATTAGT  
TTGACAACTGGGAAAGAAAGTGTCTTCAACCCCTACCCCAAGACATTCTCTGTGCGCCAGGCTGGAGTGCA  
GCCTCAACCTCCTGGGCCCCAAGTGATCCTCCACCTCAGCACACAACACCATGCCCAATTTTAAAGTGCCTTATAG  
AGACGGGGGTCTCACTTTGTTACCCAGGCTGGTCTCAAACCTCTGCGCTCAAGCAATCCTCCCACCTGGGCCCTC  
CAAAATGCTGGGTGACAGGCATGAGCGCTGTGCTGGCTCATTTTTCAGATGAGACATTTGACTGTGGCTA  
TGTAAGGAGAACATTTCTTGTCTTAGCAAAACATACTGAAGTTTTTATAGATTAATTACCACAGTGTCTGCCACTGA  
ATTTCCAGTGACTAAGTGAAAAATATAAAACATATGAATATAAAGAAAGAAAGAGACAAGTCAAATGTAGTAAA  
ATGACAACACTTGGTGACTCTAGGTGACTGGTCGACAGATGTTCAATTGTAATATCAATGTGGCTTTGCTGTGGGT  
TTGAAATTTTGCAAACTAAGAGTTGGTGGCGGGGAGAAAGGATACACCAAAAAAACTAAGTGATTATCTTTGGATG  
GAAAAATGTTTGGTAATTGCATTTTAAATGTCTTCTTGTATTTTTTAAATGTTCAATAATGTATATGTATCAG  
TTCTGTAATAAAGGGGAAAAACACTTTTCA

**FIGURE 124**

MVDLSVSPDSLKPVSILTSSLVFLMHLLLLQPGEPSSEVKVLGPEYPILALVGEEVEFPCHLWP  
QLDAQQMEIRWFRSQTFNVVHLYQEQQELPGRQMPAFRNRTKLVKDDIAYGSVVLQLHSIIPS  
DKGTYGCRFHSDNFSGEALWELEVAGLGSDPHLSLEGFKEGGIQLRLRSSGWYPKPKVQWRDH  
QGQCLPPEFEAIVWDAQDLFSLETSSVVVRAGALSNVSVSIQNLLLSQKKELVVQIADVFPVGA  
SAWKSAFVATLPLLLVLAALALGVLRKQRRSREKLRKQAEKRQEKLTAELEKLQTELDWRRAE  
GQAEWRAAQKYAVDVTLDPASAHPSLEVSEDGKSVSSRGAPPGPAPGHPQRFSEQTCALSLE  
FSAGRHYWEVHVGRRSRWFLGACLAAPRAGPARLSPAAGYWVLGLWNGCEYFVLAPHRVALT  
LRVPPRRLGVFLDYEAGELSFFNVSDGSHIFTFHDTFSGALCAYFRPRAHDGGEHPDPLTICP  
LPVRGTGVPEENDSDTWLQPYEPADPALDWW

**Important features of the protein:****Signal peptide:**

amino acids 1-34

**Transmembrane domain:**

amino acids 247-272

**N-glycosylation sites.**

amino acids 102-106, 139-143, 224-228, 464-468, 516-520

**Tyrosine kinase phosphorylation site.**

amino acids 105-114

**N-myristoylation sites.**

amino acids 129-135, 220-226, 399-405, 423-429, 480-486

**Amidation site.**

amino acids 390-394



**FIGURE 125**

TATAGTCCCAGCTACTCATGGGGCTGATGCAGGTTGAGGCAGGAGGTTTCATGAGCCCAGGAGGTTGGAGCTGTAA  
TGAGCTAGGATTTCTGCCCTCTGCACTCCTAGCTGGATGACAGAGCAAGACCCCTGTCTCAAAAAAGAAAAA  
AAAAAGAATGCATGAACCAGACATGACAGTTCCCTGGCCTCAAAGATCTTCCAAAGGAAATGATTTTTTTTAAACC  
ACCAATGCTGCAGGAAAAAGCAACATATTTAAGTTATCCAATAACACCTATCCAATAATTGTAAATCATTATCAT  
GACATGGTAGAGTTGTTTATATTTCTTTTCTTTTAGGTGAAACACCATTCAAAGTCGTAGTCAAATCTCTTTCA  
CCTAAAGAGTTGGTCCGGATACATGTCCCTAAACCTTTGGACAGGAATGATGGAACATTTTTGATGAGATATAGG  
ATGTATGAAACTGTCGATGAAGGCCCTGAAGATAGAGGTCCTTTATGGTGATGAACATGTGGCTCAGTCTCCCTAT  
ATTTTGAAAGGACCAGTGTACCATGAGTACTGTGAGTGTCCGGAAGATCCTCAGGCCCTGGCAGAAGACTCTTTCT  
TGTTCAACCAAGGAACCACAGATTGCAAAAGATTTTGCTTCTTTTCCCAGCATCAATCTCCAGCAAATGCTAAAA  
GAAGTCCCCAAAAGGTTTGGGGATGAGAGAGGTGCCATTGTTTCATTACAGATTCTCAATAACCATGTTTACCGG  
AGATCTTTAGGGAAATACACAGACTTCAAGATGTTCTCTGATGAGATTTTGTTATCATTGACAAGAAAGGTCCTT  
CTCCCAGATTTAGAATTTTATGTTAATCTTGGAGATTGGCCCTTGGAGCATCGAAAAGTCAATGGAACCCCTAGC  
CCCATACCTATCATTTTCATGGTGTGGCTCTCTGGATTCAAGAGATGTTGTCTTCCAACGTATGACATCACCCAC  
TCCATGCTTGAAGCCATGCGGGGTGTTACAAATGATCTCCTCTCTATTTCAGGGAAATACAGGGCCTTCTTGGATC  
AATAAACAGAGAGAGCTTTCTTCAGAGGTAGAGACAGCCGAGAGGAGAGGCTCCAGTTGGTACAGCTGTCCAAA  
GAAAATCCTCAGCTACTAGATGCAGGAATTACAGGATATTTCTTTTCCAAAGAGAAAAGAAAGGAGCTTGGAAAA  
GCCAAGTTGATGGGTTTCTTTGATTTCTTTAAGTACAAGTATCAAGTAAATGTGGATGGGACCGTGGCTGCTTAC  
AGATATCCATATCTCATGCTGGGCGACAGTCTGGTTTTAAAGCAGGACTCGCCATATTATGAACATTTCTACATG  
GCACTAGAACCCTTGGAAAGCATTATGTTCCAATTAAAAAGAAATCTGAGTGATTTATAGAGAAAGTTAAATGGGCT  
AAGGAAAAATGATGAAGAAGCCAAGAAGATTGCAAAAGAAGGACAGTTGATGGCTAGGGACCTACTACAGCCACAC  
AGGCTTTACTGCTACTATTACCAAGTACTGCAGAAATATGCCGAGCGCCAGTCCAGCAAACCCGAAGTACGTGAT  
GAAATGGAACCTTGTTTCCCTCAGCCAGAAGATAGCACAGCCATCTGCCAGTGCCACAGGAAAAAGCCTTCAAGAGAA  
GAACCTTGAAGTGCAGCCCAGAATCACACTCCTGTGTATCCCGGCTACACTTTAAGGAAAGATTGAATCTAAGCTGT  
GAAGGACAGTATAGAAGACTGCACCAAGTGGACTAGTTCTCCCGTGGCTTTATATATGTAGATGGATATAGCAG  
TACTGGTTGAGTATCCCTCATCTGAAATGCTTAGGACCAGGAGTGTTCAGGCTTCAGATTTTTTAAGATTTGGG  
AATATTTGCAATGTACATAATGAGGTATCTTGGGGATGAGATCCAAGTCTAAACACAAAATTCATTTATATTTTAT  
ATATACCTTGTTCACATACCTGAAGGTAATTTTATATAATATTTTAAATAATTTGTGCATGAAACAAAGTTTGT  
ATACATTTGAACCTGTCAGAAAAGCAAAGGTGTCACTATCTTAGCGACCCAAAGTGGTGGTGTGAGCACTCAAAAAGTT  
TTGGATTTTGGGGTATTTTCAAGTTTTAGATTTTTGTATGAGGAATGTTCAACCTGTATTTGAACAAGCATTACCA  
AATATCATTGAATATTAATATCTTTTGCCTAAAACTGCTATTATCAGCATCATAGTTTCTCTAAAAAGAAAAC  
TGGGGATCATAGCCGATAGAGAGACTTGCTAAAAATATAAATCAGCCTCTGCAAAACCTGTTTACATATTTATTTGGT  
TTACATATTTTATTTGTTTATTTCTATCCCTGTTCACCTTTTCTCTTCCACTTCCAATTTATGAAGAGAAAAATAT  
TTGTTTCAGGTTTCTTTTGAAGGTCTGATAAGGATATTTAAGGAAGAGAAATGACTCTGTTATTAAGGTGGCAT  
TAACAGCTTCCCTTTTGAAGGTCTGATAAGGATATTTAAGGAAGAGAAATGACTCTGTTATTAAGGTGGCAT  
GGAGACTGTGGAGGGAATATTTTTTAAAGCACTACTCATATCCTTTAAACTAAATTTTGGCCAAAGCCGAGACAA  
CATTAAAGGAGAAATTTGTACCTTAAGTTAGTAATTTCCAATCTATCTGAGTTGTATACCCATCAAAGACAATACAG  
TTATTAACATAGATGAAGGTATGCTATAGGCATCATTTCATTATCTCTATATTGAATAGGTGAAAGATAACTGTAG  
TCAGGTGAAAGGCATTCATCATTTTTAAGCTGAAAAGGGGATCCTTGAAAACACTGAAAACCTCTACAACAATCT  
TCAGGAAGCCTGCTATCTTGGGATTCACATAATAGGCAACAAAGGCAAGCATCCATTCTCCTACTCCACC  
ACTTTTCTATTTTCAGTGGGTGTCATTGCTACGATGAAGCACTTTGGAAATTTCTTTTCTTTTAGGACAGGTCGA  
GGATTTAGGACTCATAGCCTGAAAGCTCATTACATACTCCTTGTAACCATCAGTCCAAGGTTTCAGTTCACTAAAG  
TGCATGTTCTAAAAACAAGAGCTATCCTCATTTCAAATTTTAAAAATATGTACTCTGGCCGTTGTCAGTGGCTCACG  
CCTGTAATCCCAGCACTTTGGCAGGCCGAGATGGGCGGATCTTTTGAGGTCAGGAGTTTGAGACCAGCCTGGCCA  
ACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGGCATGGTGGCATTTGCTGTAAATCCCAGCT  
ACTCGGGAGGCTGAGGCAGGAGATCACTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTACACCATG  
CACTCCAGCCTGGGTGACAGAGTGAGACTCCATCTCAAAAACCTGAAAAATAAAAAATAAAAAATATGTATTTCTCTAA  
CTGAAATATTTACTTAATCTGGAAAAAATGTAACATTTTTTAAAGTGGTTACATCTATTCTTGCTGAAGAACAA  
TAAACAGAATTTTTTGAAGTAAAGCATAACCAATTTTCAAGACAGTCTAATCAATGCCAAGTATCCAAGGCAAACTC  
TAATACCCATCCATTGTGCAAAACCAAGCACGCAAGTATTAATAAGAGCAAGCTGTCTTGAGCCCATACCTA  
ATGAATTTGTGCTTAAATATTGTACATTGTGTTTGAAGGCTGTCAAACCTGGGATTATGGCAAGAAAGGTTGCC  
TAACCTACACCTTTCTGCCCTCAAATTTCCAGGTGCTAAAGGCTAATGGCATTTTAAACATCTTACATTTTAAAAA  
TTTTATATTGCTCTGCCAAACAGGCCAATAGTTAAAGCAAGTTGAGACAAACAGGCAGATTTCAGTGTGTGGA  
ACAGGAAGGATGTGCTTTAAAAAAGGTTGAATCCCTCAAAAATTTCTATAGGGAGACAGCAGCCTTAATCTACA  
TAATTTCTCATCTCGCAATTCAGCCGAGCCTTTAAAGAGTTAGTGTTAATGGCTTTCTGGTTTGAAGCAAAA  
ATGCATCTATGTGGTTGAAAGTTTGGGAGGAGATTACCAATATCTGAGGAGAAGATGGAGTGAAGGGAATTTCTT  
ACTTTTTGCTTTTATACCTTTCTATAATATTTAGATTTTTTTTACTGTAAGTATGGATCAAATTTGCAAAATAAAG  
AAAAATGCCAACCTTTAGAAAAAGCAATAAATGACAAAAAGATATAAACAGGAACCAAAATATTTATTTTTTC  
CATTTTGTCTTTTTTAAATCTATGTTTGAAGCTTTATATCTTGGGACTTATGTATATATATACCTTTTAAATAAA  
ATAAATTTTCTAAATAAAAAAGTTG

**FIGURE 126**

MVELFIFLFLLLGETPFKVVVKSLSPKELVRIHVPKPLDRNDGTFLMRYRMYETVDEGLKIEVL  
YGDEHVAQSPYILKGPVYHEYCECPEDPQAWQKTLSCPTKEPQIAKDFASFPSINLQQMLKEV  
PKRFGDERGAIVHYTILNNHVYRRSLGKYTDFKMFSDEILLSLTRKVLLPDLEFYVNLGDWPL  
EHRKVNGTPSPIPIISWCGSLDSRDVVLPTYDITHSMLEAMRGVTNDLLSIQGNTGPSWINKT  
ERAFFRGRDSREERLQLVQLSKENPQLLDAGITGYFFFQEKELGKAKLMGFFDFFKYKYQV  
NVDGTVAAYRYPYMLGDSLVLKQDSPYYEHFYMALEPWKHYVPIKRNLSDLLEKVKWAKEND  
EEAKKIAKEGQLMARDLLQPHRLYCYYYQVLQKYAERQSSKPEVRDGMELVPQPEDSTAICQC  
HRKKPSREEL

**Important features of the protein:****Signal peptide:**

amino acids 1-16

**N-glycosylation sites.**

amino acids 250-254, 363-367

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 444-448

**N-myristoylation site.**

amino acids 208-214, 319-325, 388-394

**Endoplasmic reticulum targeting sequence.**

amino acids 448-453

**Mitochondrial energy transfer proteins signature.**

amino acids 25-34

## FIGURE 127

AGCGCTCGGAGGGAGCCGGAGCGCTTCTCCCGAGTTGGTGATAGATTGGTGGTTCATCCAACT  
 GCAGAAATGAATGAGCAGTGAAAAGCAGCAGAGCCGATGGGTTCATGAGGATGTAAGTGCCTTT  
 GAAGGCTTCCACACCCTCTACTCCAGGAATCATGAATAAACTGGAGGATAAGCAGGACCAGAT  
 GATACCATGAAGAGAAGTTTACAGGCCCTCTATTGCCAACTGTTAAGTTTCCTGCTGATCTTG  
 GCACTGACCGAAGCGCTGGCATTTCGCATCCAGGAACCATCTCCAGGGAATCTCTTCAGGTC  
 CTCCTTCAGGCACTCCCCCGGGAACCATGGTGACAGCACCCACAGCTCTACCAGACATACT  
 TCTGTGGTGATGCTGACCCCCAATCCCGATGGACCCCCCTCACAGGCTGCAGCTCCCATGGCA  
 ACACTGACACCCCCGTGCAGAGGGGCAACCCTCCTACGCACACCATCTCCACCATCGCTGCGACA  
 GTAACCGCCCCCTATTCTGAAAGCTCCCTGTCCACAGGGCCCCGCTCCAGCAGCCATGGCAACC  
 ACATCCTCCAAGCCAGAGGGCGCCCTCGAGGGCAGGCTGCCCCCACCATCCTGCTGACAAAG  
 CCACCGGGGGCCACCAGCGCCCCACCACAGCGCCCCCGCACTACCACACGCAGGCCCCCCC  
 AGGCCCCCAGGCTCTTCCCGAAAAGGGGCTGGTAATTCATCACGCCCTGTCCC GCCTGCACCT  
 GGTGGCCACTCCAGGAGTAAAGAAGGACAGCGAGGACGAAATCCAAGCTCCACACCTCTGGGG  
 CAGAAGCGGCCCTGGGGAAAATCTTTCAGATCTACAAGGGCAACTTCACAGGGTCTGTGGAA  
 CCAGAGCCCTCTACCCTCACCCCCAGGACCCCCACTCTGGGGCTACTCCTCTTCACCACAGCCC  
 CAGACAGTGGCTGCGACCAGTGGCCAGCAATACCTCATGGGCACCCACCACCACCTCCCTG  
 GGGCCTGCAAAGGACAAGCCAGGCCCTTCGCAGAGCAGCCCAGGGGGGTGGTTCTACCTTCACC  
 AGCCAAGGAGGGACACCAGATGCCACAGCAGCCTCAGGTGCCCTGTGAGTCCACAAGCTGCC  
 CCAGTGCCTTCTCAGCGCCCCCACCACGGTGACCCACAGGATGGCCCCAGCCATAGTGACTCT  
 TGGCTTACTGTTACCCCTGGCACCAGCAGACCTCTGTCTACCACTCTGGGGTCTTCACGGCT  
 GCCACGGGGCCCCACCCAGCTGCCTTCGATACCAGTGCTCTCAGCCCTTCCCAGGGGATTCTCT  
 CAGGGAGCATCCACAACCCCAACAAGCTCCAACCCATCCCTCCAGGGTCTCAGAAAGCACTATT  
 TCTGGAGCCAAGGAGGAGACTGTGGCCACCCTCACCATGACCGACCGGGTGCCAGTCTCTC  
 TCCACAGTGGTATCCACAGCCACAGGCAATTTCTCAACCGCTGGTCCCCGCCGGGACCTTG  
 AAGCCTGGGACAGCAGGGAACATCTCCCATGTGGCCGAGGGGGACAAACCGCAGCACAGAGCC  
 ACCATCTGCCTGAGCAAGATGGATATCGCCTGGGTGATCCTGGCCATCAGCGTGCCCATCTCC  
 TCCTGCTCTGTCTCTGCTGACGGTGTGCTGCATGAAGAGGAAGAAGAAGACCGCCAACCCGGAG  
 AACAACTGAGCTACTGGAACAACACCATCACCATGGACTACTTCAACAGGCATGCTGTGGAG  
 CTGCCCAGGGAGATCCAGTCCCTTGAAACCTCTGAGGACCAGCTCTCAGAGCCCCGCTCCCCA  
 GCCAATGGCGACTATAGAGACACTGGGATGGTCCTTGTTAACCCCTTCTGTCAAGAAACACTG  
 TTTGTGGGAAACGATCAAGTATCTGAGATCTAACTACAGCAGGCATCACTTTGCCATTCCGTA  
 TTTTTCGTCTCTAAATTATAAATATACAAATATATATATTATAAATATAACCTTGTGTAACCC  
 TGACTTAATGAGAAACATTTTCAGCTTTTTTTTCTATGAATTGTCAACATCTTTTTTTACAAGT  
 GTGGTTTTAAAAAAAAAAAAAACTTTACAGAATGATCTGTGGCTTTATAAAATAAAGGTATTTCT  
 AAGCAAAAAAAAAAAAAAAAAA

**FIGURE 128**

MKRSLQALYCQLLSFLLILALTEALAFAIQEPPRESLQVLPSGTPPGTMVTAPHSSSTRHTSV  
 VMLTPNPDGPPSQAAAPMATLTPRAEGHPPTHISTIAATVTAPYSESSLSTGPAPAAMATTS  
 SKPEGRPRGQAAPTILLTKPPGATS RPTTAPPRTTTRRPPRPPGSSSRKGAGNSSRPVPPAPGG  
 HSRKEGQRGRNPSSTPLGQKRPLGKIFQIYKGNFTGSVEPEPSTLTPRTPLWGYSSSPQPQT  
 VAATTVPNTSWAPTTTSLGPAKDKPGLRRAAQGGGSTFTSQGGTPDATAASGAPVSPQAAPV  
 PSQRPHHGDPQDGPSHSDSWLTVTPGTSRPLSTSSGVFTAATGPTPAAFDTSVSAPSQGIPOG  
 ASTTPQAPTHPSRVSESTISGAKEETVATLTMTDRVPSPLSTVVSTATGNFLNRLVPAGTWKP  
 GTAGNISHVAEGDKPQHRATICLSKMDIAWVILAI SVPISSCSVLLTVCCMKRKKKTANPENN  
 LSYWNNTITMDYFNRHAVELPREIQSLETSEDQLSEPRSPANGDYRDTGMVLVNPFCQETLFV  
 GNDQVSEI

**Important features of the protein:****Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 469-487

**N-glycosylation sites.**

amino acids 178-182, 223-227, 261-265, 446-450, 504-508, 509-513

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 495-499

**N-myristoylation sites.**

amino acids 44-50, 48-54, 175-181, 222-228, 279-285, 286-292,  
288-294, 296-302, 351-357, 374-380, 427-433, 442-448

**TonB-dependent receptor proteins signature 1.**

amino acids 1-44

**FIGURE 129**

AGGCGAGGCGCGCGCGCGCTGCACACACGCACGAGCT**ATG**GGGGTGCCATGTTGCCACCAG  
 CTGCCACGTGGCCTGGCTTTTTGGTGCTGATCTCTGGATGCTGGGGCCAGGTGAACCGGCTGCC  
 CTTCTTTCACCAACCACTTCTTTGATACATACCTGCTGATCAGCGAGGACACGCCTGTGGGTTCT  
 TTCTGTGACCCAGTTGCTGGCCCCAAGACATGGACAATGACCCCCCTGGTGTTTGGCGTGTCTGG  
 GGAGGAGGCCTCTCGCTTCTTTGCAGTGGAGCCTGACACTGGCGTGGTGTTGGCTCCGGCAGCC  
 ACTGGACAGAGAGACCAAGTCAGAGTTTACCCTGGAGTTCTCTGTGACGACCACCAGGGGGT  
 GATCACACGGAAGGTGAACATCCAGGTCGGGGATGTGAATGACAACGCGCCACATTTACAA  
 TCAGCCCTACAGCGTCCGCATCCCTGAGAATACACCAAGTGGGGACGCCCATCTTCATCGTGAA  
 TGCCACAGACCCCGACTTGGGGGCGAGGGGACGCGTCCTCTACTCCTTCCAGCCCCCTCCCA  
 ATTCTTCGCCATTGACAGCGCCCGCGGTATCGTCAACAGTGATCCGGGAGCTGGACTACGAGAC  
 CACACAGGCCTACCAGCTCACGGTCAACGCCACAGATCAAGACAAGACCAGGCCTCTGTCCAC  
 CCTGGCCAACCTTGGCCATCATCATCACAGATGTCCAGGACATGGACCCCATCTTCATCAACCT  
 GCCTTACAGCACCAACATCTACGAGCATTCTCCTCCGGGCACGACGGTGCGCATCATCACCGC  
 CATAGACCAGGATAAAGGACGTCCCCGGGGCATTGGCTACACCATCGTTTTCAGGGAATACCAA  
 CAGCATCTTTGCCCTGGACTACATCAGCGGAGTGCTGACCTTGAATGGCCTGCTGGACCGGGA  
 GAACCCCTGTACAGCCATGGCTTCATCCTGACTGTGAAGGGCACGGAGCTGAACGATGACCG  
 CACCCCATCTGACGCTACAGTACCACGACCTTCAATATCCTGGTTATTGACATCAATGACAA  
 TGCCCCGGAGTTCAACAGCTCCGAGTACAGCGTGGCCATCACTGAGCTGGCACAGGTCGGCTT  
 TGCCCTTCCACTCTTCATCCAGGTGGTGACAAAGGATGAGAATTTGGGCCTGAACAGCATGTT  
 TGAGGTGTACTTGTTGGGGAACAACCTCCACCACTTCATCATCTCCCCGACCTCCGTCCAGGG  
 GAAGGCGGACATTTCGTATTCGGGTGGCCATCCCACTGGACTACGAGACCGTGGACCGCTACGA  
 CTTTGATCTCTTTGCCAATGAGAGTGTGCCTGACCATGTGGGCTATGCCAAGGTGAAGATCAC  
 TCTCATCAATGAAAATGACAACCGGCCCATCTTCAGCCAGCCACTGTACAACATCAGCCTGTA  
 CGAGAACGTCACCGTGGGGACCTCTGTGCTGACAGTCCTGGTGAGTCCCCGCTTCACTGCAGG  
 GCCACTGAGCTCTCCAGGGCCGACTGTGGTGAGGCACCCAGAGGGATTTTGTCCAAGGGACCT  
 CAGCAATCAGGGAAGGAGGCACCCCCAAATCCCTGAGCTGTGTTTGTGGTGAT**TAA**ATAAA  
 GTTTTTTGACTCTTCAGGAAGGGGCTCCCTTGACCTAGGTTGCAATATGGAAAAGGAGCCAAC  
 CTGAGGGGTGACGAGACTGAGCTGAGGACACTGGTTTTCTGCCTTTCCTGAGAGAGACTCAG  
 TGAGGGTGGGCTGGGAGCCCTGGAAGCCCCCTCAAATGGGTGGGAAGGTGCCAGCCATCCTTG  
 AGAAGGGCAACCCCTCTCCATGTGAGCACAGGCACAGAGAGGGGCAGGCGCTTGGAGGGTACC  
 GGGGCACCCCCAGCTGCCCATGGCTGGACTTGCCCTTTGACAAGGGGCCCTCCAGTGTCATT  
 TGTATCTGTCAGTACTCTTGTTTGCAAGGGACAGAAACCCCTTAAGTAGTTCAAGCAAAAAGG  
 ATTGGCTCATGTAACCTAAAAGTATAAGTGATTTACGGCCGGGCTCGGTGGCTCACGCCTGTC  
 ATCCAACACCTTGAGAAAGCCGAGTGGGCGGATCACTTGAGGTGCGGAGTTTGAGACCAGG  
 TGGCCAACATGGCAAAACCCCGTCTCTACTAAAATACAAAATTAGCCGGGTGTGGTGGCAC  
 ACGCCTGTAGTCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGCGG  
 AGGTTGCAGTGAGCCGAGATTGTGTCACTGCCCTCCAGCCTGGGCGACAGAGCCAGATTCTGT  
 CTC

**FIGURE 130**

MGCHVATSCHVAWLLVLISGCWGQVNRLPFFTNNHFFDTYLLISEDTPVGSSVTQLLAQDMDND  
 PLVFGVSGEEASRFFFAVEPDTGVVWLRQPLDRETKSEFTVEFSVSDHQGVITRKVNIQVGDVN  
 DNAPTFHNQPYSVRIPENTPVGTPIFIVNATDPDLGAGGSVLYSFQPPSQFFAIDSARGIVTV  
 IRELDYETQAYQLTVNATDQDKTRPLSTLANLAIITDVQMDPIFINLPYSTNIYEHSPPG  
 TTVRIITAIDQDKGRPRGIGYTIVSGNTNSIFALDYISGVLTNLGLLDRENPLYSHGFILTVK  
 GTELNDDRTPSDATVTTTTFNILVIDINDNAPEFNSSEYSVAITELAQVGFALPLFIQVVDKDE  
 NLGLNSMFEVYLVGNNSHHFIIISPTSVQGKADIRIRVAIPLDYETVDRYDFDLFANESVPDHV  
 GYAKVKITLINENDNRPIFSQPLYNISLYENVTVGTSVLTVLVSPRFTAGPLSSPGPTVVRHP  
 EGFCPRDLSNQRRHPQIPELCLLVY

**Important features of the protein:****Signal peptide:**

amino acids 1-23

**Transmembrane domain:**

amino acids 355-374

**N-glycosylation sites.**

amino acids 155-159, 206-210, 349-353, 393-397, 434-438, 466-470,  
 472-476

**N-myristoylation sites.**

amino acids 2-8, 49-55, 162-168, 270-276, 278-284, 316-322

**Amidation site.**

amino acids 515-519

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 11-22

**Leucine zipper pattern.**

amino acids 298-320

**PTS HPR component serine phosphorylation site signature.**

amino acids 377-393

**Cadherins extracellular repeated domain signature.**

amino acids 120-131, 336-347

**Cadherins extracellular**

amino acids 120-144, 336-360



## **FIGURE 132**

MLMFAVIVASSGLLLMIERGILAEKPLPLHPPGREGTAWRGKAPKPGGLSLRAGDADLQVRQ  
DVRNRTLRAVCGQPGMPRDPWDLVVGQRRTLLRHILVSDRYRFLYCYVPKVACSNWKRVMKVL  
AGVLDSVDVRLKMDHRSDLVFLADLRPEEIIRYRLQHYFKFLFVREPLERLLSAYRNKFGEIRE  
YQORYGAEIVRRYRAGAGPSPAGDDVTFPEFLRYLVDEDPERMNEHWMPVYHLCQPCAVHYDF  
VGSYERLEADANQVLEWVRAPPHVRFPARQAWYRPASPESLHYHLC SAPRALLQDVLPKYILD  
FSLFAYPLPNVTKEACQQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-23

#### **N-glycosylation sites.**

amino acids 67-71, 325-329

#### **Tyrosine kinase phosphorylation sites.**

amino acids 152-159, 183-183

#### **N-myristoylation sites.**

amino acids 89-95, 128-134



**FIGURE 133**

CGGCAGTTCTGGCCCCCTGCAGCTGGAGGTACCCTGAGTTCTGAGGGTCGTAGTGCTGTTTCTG  
 GTATTCTCATCGCGGTACCTCTACCGGTGTGGACAAGTAAAGTTTGAATCAGCTTCTCCATG  
 GCCTGGGCACCAGTTCCCGGCTGAGCCATTTTCCTTTTGGCTAAAAGTCCCCGCCAGAGGCC  
 AATTCGTCGCGGCGGCGGTGGAGATCGCAGGTGCTCAGGCTTGCAG**ATG**GGTCAAGGGTTGT  
 GGAGAGTGGTCAGAAACCAGCAGCTGCAACAAGAAGGCTACAGTGAGCAAGGCTACCTCACCA  
 GAGAGCAGAGCAGGAGAATGGATGCGAGCAACATTTCTAACACCAATCATCGTAAACAAGTCC  
 AAGGAGGCATTGACATATATCATCTTTTGAAGGCAAGGAAATCGAAAGAACAGGAAGGATTCA  
 TTAATTTGGAAATGTTGCCTCCTGAGCTAAGCTTTACCATCTTGTCCTACCTGAATGCAACTG  
 ACCTTTGCTTGGCTTCATGTGTTTGGCAGGACCTTGCGAATGATGAACTTCTCTGGCAAGGGT  
 TGTGCAAATCCACTTGGGGTCACTGTTCCATATACAATAAGAACCCACCTTTAGGATTTTCTT  
 TTAGAAAATTGTATATGCAGCTGGATGAAGGCAGCCTCACCTTTAATGCCAACCCAGATGAGG  
 GAGTGAACACTTTTATGTCCAAGGGTATCCTGGATGATTCGCCAAAGGAAATAGCAAAGTTTA  
 TCTTCTGTACAAGAACACTAAATTGGAAAAAACTGAGAATCTATCTTGATGAAAGGAGAGATG  
 TCTTGGATGACCTTGTAACATTGCATAATTTTAGAAATCAGTTCTTGCCAAATGCACTGAGAG  
 AATTTTTTCGTATATCCATGCCCCCTGAAGAGCGTGGAGAGTATCTTGAACTCTTATAACAA  
 AGTTCTCACATAGATTCTGTGCTTGCAACCCTGATTTAATGCGAGAACTTGGCCTTAGTCCTG  
 ATGCTGTCTATGTACTGTGCTACTCTTTGATTCTACTTTCCATTGACCTCACTAGCCCTCATG  
 TGAAGAATAAAATGTCAAAAAGGGAATTTATTCGAAATACCCGTCGCGCTGCTCAAAATATTA  
 GTGAAGATTTTGTAGGGCATCTTTATGACAATATCTACCTTATTGGCCATGTGGCTGCA**TAAA**  
 AAGCACAATTGCTAGGACTTCAGTTTTTACTTCAGACTAAAGCTACCCAAGGACTTAGCAGAT  
 ATGGGGGTTACATCAGTGCTGGTCATTGTAGCCTGAGTATACAATCAAGCTTCAGTGTGCAAC  
 CTTTTTTTCTTTTGCCATTTTCTATTTTAGTAATTTTCTTGGGGAATAAATAATTTTGCAGA  
 ATTTTTCCTAATTTTGTATTATCACGTTTTGCACAAAGCAGAGCCACTGTCTAACACAGCTGTT  
 AACGAATGATAAACTGACATTATACTCTAAAAGATGGTGTATTTGTGCATTAGATTTGCCTGA  
 AAACTTTATCCATTTCCATTCTTTATACAAATACCATGTAATGTGTACATATTTAACTAAAG  
 AGATTTATAGTCATAATTATTTTATTGTAAAGATTTTAACTAAAGTTTTTCCTTTTCTCTC

**FIGURE 134**

MGQGLWRVVRNQQLQQEGYSEQGYLTREQSRRMDASNISNTNHRKQVQGGIDIYHLLKARKSK  
EQEGFINLEMLPPELSFTILSYLNATDLCLASCWQDLANDELLWQGLCKSTWGHCSIYNKNP  
PLGFSFRKLYMQLDEGSLTFNANPDEGVNYFMSKGILDDSPKEIAKFIFCTRNLNWKKLRIYL  
DERRDVLDDLVTLHNFRNQFLPNALREFFRHIHAPEEERGEYLETLITKFSHRFCACNPDLMRE  
LGLSPDAVYVLCYSLILLSIDLTSPHVKNKMSKREFIRNTRRAAQNISEDFVGHLYDNIYLI  
GVAA

**Important features of the protein:****Transmembrane domain:**

amino acids 253-272

**N-glycosylation sites.**

amino acids 37-41, 87-91, 298-302

**N-myristoylation site.**

amino acids 110-116

**FIGURE 135**

GGCACGAGGGAGCCTCCGTTAGGGGGTGGGAAAGGACTTTGCCATAGGTCGCTGAGGCCACCA  
TCTGCTCTCTTACTGGCCAAGGGCGTAAAAAGATAGTCTTCCCATTAGCTAGAGAGCAAACCC  
CAGAAAGCCTATTGGCTGCGCCGTCCGCGGGCCTTGGTCCGCTTTGAAGGCGGGCTGCGGCTG  
CGAGAGGAGGGCGGGCGGGAGGCTAGCTGTTGTCTGCTGGTTGCTCGGAGGCACGTGTGCAGTCC  
CGGAAGCGGCGAGGGGAAACTGCTCCGCGCGCGCCGCGGGAGGAGGAACCGCCCGGTCCTTTT  
GGGTCCGGGCCCCGGCCGGGCCATGGATTCAATGCCTGAGCCCGCGTCCCGCTGTCTTCTGCTT  
CTTCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCCGGCCCCGGAGCTGGGCCCCGAGCCAGGCC  
GGAGCTGAGGAGAACGACTGGGTTTCGCTGCCAGCAAATGCGAAGTGTGTAAATATGTTGCT  
GTGGAGCTGAAGTCAGCCTTTGAGGAAACCGGCAAGACCAAGGAGGTGATTGGCACGGGCTAT  
GGCATCCTGGACCAGAAGGCCTCTGGAGTCAAATACACCAAGTCGGACTTGCGGTTAATCGAA  
GTCACTGAGACCATTGTGAAGAGGCTCCTGGATTATAGCCTGCACAAGGAGAGGACCGGCAGC  
AATCGATTTGCCAAGGGCATGTCAGAGACCTTTGAGACATTACACAACCTGGTACACAAAGGG  
GTCAAGGTGGTGATGGACATCCCCTATGAGCTGTGGAACGAGACTTCTGCAGAGGTGGCTGAC  
CTCAAGAAGCAGTGTGATGTGCTGGTGGAAGAGTTTGAGGAGGTGATCGAGGACTGGTACAGG  
AACCACCAGGAGGAAGACCTGACTGAATTCCTCTGCGCCAACCACGTGCTGAAGGGAAAAGAC  
ACCAGTTGCCTGGCAGAGCAGTGGTCCGGCAAGAAGGGAGACACAGCTGCCCTGGGAGGGAAG  
AAGTCCAAGAAGAAGAGCAGCAGGGCCAAGGCAGCAGGCGGCAGGAGTAGCAGCAGCAAACAA  
AGGAAGGAGCTGGGTGGCCTTGAGGGAGACCCAGCCCCGAGGAGGATGAGGGCATCCAGAAG  
GCATCCCCTCTCACACACAGCCCCCTGATGAGCTCTTGAAGCCCACCCAGCATCCTCTGTCTTG  
AGACCCCTGATTTTGAAGCTGAGGAGTCAGGGGCATGGCTCTGGCAGGCCGGGATGGCCCCGC  
AGCCTTCAGCCCCTCCTTGCCCTTGCTGTGCCCTCTTCTGCCAAGGAAAGACACAAGCCCCAG  
GAAGAACTCAGAGCCGTCATGGGTAGCCACGCCGTCTTTCCCCTCCCCAAGTGTTTCTCTC  
CTGACCCAGGGTTTCAGGCAGGCCTTGTGGTTTCAGGACTGCAAGGACTCCAGTGTGAACCTCAG  
GAGGGGCAGGTGTCAGAACTGGGCACCAGGACTGGAGCCCCCTCCGGAGACCAAACCTACCCAT  
CCCTCAGTCCTCCCCAACAGGGTACTAGGACTGCAGCCCCCTGTAGCTCCTCTCTGCTTACCC  
CTCCTGTGGACACCTTGCACTCTGCCTGGCCCTTCCCAGAGCCCCAAAGAGTAAAAATGTTCTG  
GTTCTGATTTCTGAAAAAAAAAAAAAAAAAAAAATTCCT

**FIGURE 136**

MDSMPEPASRCLLLLPLLLLLLLLLLPAPELGPSQAGAEENDWVRLPSKCEVCKYVAVELKSAF  
 EETGKTKEVIGTGYGILDQKASGVKYTKSDLRLIEVTETICKRLLDYSLHKERTGSGNRFAKGM  
 SETFETLHNLVHKGVKVVM DIPYELWNETS AEVADLKKQCDVLVEEFEEVIEDWYRNHQEEDL  
 TEFLCANHVLKGKDTSCLAEQWSGKKGDTAALGGKKSKKKSSRAKAAGGRSSSSKQRKELGGL  
 EGDPSPEEDEGIQKASPLTHSPPEL

**Important features of the protein:****Signal peptide:**

amino acids 1-26

**N-glycosylation site.**

amino acids 153-157

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 227-231, 228-232

**Tyrosine kinase phosphorylation site.**

amino acids 142-150

**N-myristoylation sites.**

amino acids 36-42, 74-80, 86-92, 125-131, 222-228, 237-243,  
 250-256, 263-269

**Amidation sites.**

amino acids 212-216, 222-226

**ATP/GTP-binding site motif A (P-loop).**

amino acids 62-70

## FIGURE 137

CACGCTCCCGCTGCCAGCCCGGACCGGATCTTAATCAGTCACCTATGAAAACATCATTAGCT  
CCACAGCAATGAGTCCTCCACTGCTGAAGCTTTGGCGCTGTGCTTAGTACCATGGCAATGATCT  
CAAACCTGGATGTCCCAAACCTCTCCCATCCTTGGTGGGACTGAACACCACGAGGCTGTGCACTC  
CGGATACCTTAACTCAGATTAGTCCTAAAGAAGGGTGGCAGGTGTACAGCTCAGCTCAGGATC  
CTGATGGGCGGTGCATTTGCACAGTTGTTGCTCCAGAACAAAACCTGTGTTCCCGGGATGCCA  
AAAGCAGGCAACTTCGCCAACTACTGGAAAAGGTTCAGAACATGTCCCAGTCTATTGAAGTCT  
TAAACTTGAGAACTCAGAGAGATTTCGAATATGTTTTAAAAATGGAACCCAAATGAAAGGGC  
TGAAGGCAAAATTTCCGCAGATTGAAGATGATCGAAAGACACTTATGACCAAGCATTTTTCAGG  
AGTTGAAAGAGAAAAATGGACGAGCTCCTGCCTTTGATCCCCGTGCTGGAACAGTACAAAACAG  
ATGCTAAGTTAATCACCCAGTTCAAGGAGGAAATAAGGAATCTGTCTGCTGTCTCCTCACTGGTA  
TTCAGGAGGAAATTGGTGCCTATGACTACGAGGAACTACACCAAAGAGTGCTGAGCTTGGAAA  
CAAGACTTCGTGACTGCATGAAAAAGCTAACATGTGGCAAACCTGATGAAAATCACAGGCCCG  
TTACAGTCAAGACATCTGGAACCCGATTTGGTGGCTTGGATGACAGACCCTTTAGCATCTGAGA  
AAAACAACAGAGTCTGGTACATGGACAGTTATACTAACATAAAATTGTTGTTGAATACAAAT  
CAATTGCAGACTTTGTGCTGAGTGGGGCTGAATCAAGGACATACAACCTTCCCTTCAAGTGGGCAG  
GAACTAACCATGTTGTCTACAATGGCTCACTCTATTTTAAACAAGTATCAGAGTAATATCATCA  
TCAAATACAGCTTTGATATGGGGAGAGTGCTTTGCCAACGAAGCCTGGAGTATGCTGGTTTTCT  
ATAATGTTTACCCCTACACATGGGGTGGATTCTCTGACATCGACCTAATGGCTGATGAAATCG  
GGCTGTGGGCTGTGTATGCAACTAACCAGAATGCAGGCAATATTGTCATCAGCCAACTTAACC  
AAGATACCTTGAGAGTGATGAAGAGCTGGAGCACTGGCTACCCCAAGAGAAGTGAGGGGAAT  
CTTTCATGATCTGTGGGACACTGTATGTACCAACTCCCACTTAACTGGAGCCAAAGGTGTATT  
ATTCCATATTCCACCAAACCTCCACATATGAGTACACAGACATTCCCTTCCATAACCAATACT  
TTCACATATCCATGCTTGACTACAATGCAAGAGATCGAGCTCTCTATGCCTGGAACAATGGCC  
ACCAGGTGCTGTTCAATGTCACCCTTTTCCATATCATCAAGACAGAGGATGACACATAGGCAA  
ATGTGACATGTTTTTCATTGATTTAAACAGTGTGATTTGTGATAAACTCTATAAGACCCCTTCC  
GTTTTTTTTCTTCACTATTATTTTTTCATCATTTCTCCAAAGCAAAGCATTTTTATTGTAAAGTT  
GGTGTFTTCAAAAACATAGCTGAGCTTGTCTAACTTACCATGTTGGAAACACATCTTAACTTCT  
AAATTTACAAGGCCTATCATGTCCTTGTGATGAAAAGCACTAAAAAAAAAAAAAGAGTTTAAGT  
GGCTAAAGTCATAGTTTTTGCAAGAGATTAATGATCTGCCTTATATTAGAGTCAGAGACTAATG  
GTGGCTTAAATGCACGAATGTCTTTTTTTTTTAAACTGTCATTTTTTTACTGTCTTTTGCTCCA  
TCTCAGGAAATATTTTGGTAGGAATTAGGAGAACAAAAAGCACTTTTATCCCATTTATTTCTT  
TAAAAAATGTAAGGATTTCATTTATATTGAAAAATAATATTAATCATTTTGTGTTAACACAA  
TTCTCTGATGCGGTGCTGTACAGTCATTTTTTAAATCTCTTGCTAACATTTTATTGGCAGTATG  
TATTTCTACCATTTGTAACCACCATTTGTGCTATTGTATCTCTTCACTTCTGTGAAAGTAATATT  
TTTTTATAAANACACTGNAATTTTTAAA

**FIGURE 138**

MSPPLLKLGAVLSTMAMISNWMSQTLPSLVGLNTTRLSTPDTLTQISPKEGWQVYSSAQDPDG  
RCICTVVAPEQNLCSDAKSRQLRQLLEKVNMSQSIEVLNLRRTQRDFQYVLKMETQMKGLKA  
KFRQIEDDRKTLMTKHFQELKEKMDSELLPLIPVLEQYKTDAKLITQFKEEIRNLSAVLTG IQE  
EIGAYDYEELHQRVLSLETRLRDCMKKLTGKLMKITGPVTVKTSGTRFGAWMTDPLASEKNN  
RVWYMDSYTNKIVREYKSIADVFSGAESRTYNLPFKWAGTNHVYNGSLYFNKYQSNII IKY  
SFDMGRVLAQRSLEYAGFHNVPYPTWGGFSDIDLMADEIGLWAVYATNQNAGNIVISQLNQDT  
LEVMSKSWSTGYPKRSAGESFMICGTLYVTNSHLTGAKVYYSYSTKTSTYEYTDIPFHNQYFHI  
SMLDYNARDRALYAWNNGHQVLFNVTLFHI IKTEDDT

**Important features of the protein:****Signal peptide:**

amino acids 1-16

**N-glycosylation sites.**

amino acids 33-37, 95-99, 179-183, 299-303, 465-469

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 215-219

**Tyrosine kinase phosphorylation site.**

amino acids 106-114

**N-myristoylation sites.**

amino acids 9-15, 31-37, 235-241, 239-245

**FIGURE 139**

GAAGCAGTGCAGAGAGGAGAGCGGAGCGGAGCTGCCGCTGAGCAAAGGCCCTTCACCAATGGCCCG  
AGTCCCCCGGCTGCTGCTCCGTCTGGGCCCCGCTGCCTCCACTGCCTGTATAGCTGCCACTGGA  
GGAAATGCCCCAGAGAGAGGATGCAAACCAGCAAGTGCGACTGTATCTGGTTTGGCCTGCTCT  
TCCTCACCTTCCTCCTTTCCCTGAGCTGGCTGTACATCGGGCTCGTCCTTCTCAATGACCTGC  
ACAACCTTCAATGAATTCTCTTCCGCCGCTGGGGACACTGGATGGACTGGTCCCTGGCATTCC  
TGCTGGTCATCTCTCTACTGGTCACATATGCATCCTTGCTATTGGTCCTGGCCCTGCTCCTGC  
GGCTTTGTAGACAGCCCCTGCATCTGCACAGCCTCCACAAGGTGCTGCTGCTCCTCATTATGC  
TGCTTGTGGCGGCTGGCCTTGTGGGACTGGACATCCAATGGCAGCAGGAGTGGCATAGCTTGC  
GTGTGTCACTGCAGGCCACAGCCCCATTCTTTCATATTGGAGCAGCCGCTGGAATTGCCCTCC  
TGGCCTGGCCTGTGGCTGATACCTTCTACCGTATCCACCGAAGAGGTCCCAAGATTCTGCTAC  
TGCTCCTATTTTTTTGGAGTTGTCCTGGTCATCTACTTGGCCCCCCTATGCATCTCCTCACCTC  
GCATCATGGAACCCAGAGACTTACCACCCAAGCCTGGGCTGGTGGGACACCGAGGGGGCCCCCA  
TGCTGGCTCCCAGAAACACCCTGATGTCCTTGCGGAAGACAGCTGAATGCGGAGCTACTGTGT  
TTGAGACTGATGTGATGGTCAGCTCCGATGGGGTCCCCTTCTCATGCATGATGAGCACCTCA  
GCAGGACCACGAATGTAGCCTCTGTATTCCCAACCCGAATCACAGCCCACAGCAGTGACTTCT  
CCTGGACTGAACTGAAGAGACTCAATGCTGGATCCTGGTTCTTAGAGAGGCGACCCTTCTGGG  
GGGCCAAACCGCTGGCAGGCCCTGATCAGAAAGAGGCTGAGAGTCAGACGGTACCAGCATTAG  
AAGAGCTATTGGAGGAAGCTGCAGCCCTCAACCTTTCCATCATGTTTCGACTTGCGCCGACCCC  
CACAGAACCACACATACTATGACACTTTTGTGATCCAGACATTGGAGACTGTGCTGAATGCAA  
GGGTGCCCCAAGCCATGGTCTTTTGGCTACCAGATGAAGATCGGGCTAATGTCCAACGACGGG  
CACCTGGAATGCGCCAGATATATGGACGTCAGGGAGGCAACAGAACGGAGAGGCCCCAGTTTC  
TTAACCTCCCCTATCAAGATCTGCCACTATTGGATATCAAGGCATTGCATAAGGATAATGTCT  
CGGTGAACCTATTTGTAGTGAACAAGCCCTGGCTCTTCTCTCTGCTTTGGTGTGCAGGGGTGG  
ATTCGGTCAACCACCAACGACTGCCAGCTGCTGCAGCAGATGCGTTACCCTATCTGGCTTATTA  
CCCCTCAAACCTACCTAATCATATGGGTCAATTACCAATTGTGTTTTCCACCATGCTGCTTTTTGT  
GGACCTTCCTCCTCCAAAGGAGATTTGTTAAGAAGAGAGGGGAAAACCTGGCTTAGAAACAGCAG  
TGCTGCTGACAAGGATCAACAATTTTCATGATGGAGTGAATGCCCTGCCCTGCTTCCCCACCCA  
AGCCAGTCTACATTGCCCAAACAGCAAGGGTTGGAGAGTGGCTTAAGTGGAATGCTTCAGGGG  
TGGTGGGTTGCAAGTGGGGGGAGCTTTGCCAACAGGAGGTTTTGAACCATGAGGGGCCCTCTGC  
CCAGGTGATGGGCATTCCCTAAGCTGCTATGGAATCTGCTCCCTTTGGGGTTTTGACCTGAGA  
TGTTTGGGAAGAGAGTGAGTAATGAGAAGTTTCTCCTCAAATGAACTAGAACAGAGGAAGTA  
AAAGGGGAGATTGCTCGGA

**FIGURE 140**

MAESPGCCSVWARCLHCLYSCHWRKCPRERMQTSKDCIWFGLLFLTFLLSLSWLYIGLVLLN  
 DLHNFNEFLFRRWGHWMDWSLAFLLVISLLVITYASLLLVLALLRLCRQPLHLHSLHKVLLLL  
 IMLLVAAGLVGLDIQWQQEWHSLRVSLQATAPFLHIGAAAGIALLAWPVADTFYRIHRRGPKI  
 LLLLLFFGVVLVIYLAFLCISSPCIMEPRDLPPKPGLVGHRGAPMLAPENTLMSLRKTAECGA  
 TVFETDVMVSSDGVPFLMHDEHLSRTTNVASVFPTRIAHSSDFSWTELKRLNAGSWFLERP  
 FWGAKPLAGPDQKEAESQTPALEELLEEEAAALNLSIMFDLRRPPQNHTYYDTFVIQTLETVL  
 NARVPQAMVFWLPDEDNANVQRRAPGMRQIYGRQGGNRTERPQFLNLPYQDLPLLDIKALHKD  
 NVSVNLFVVNKPWLFSLLWCAGVDSVTTNDCQLLQQMRYPIWLITPQTYLIIWVITNCVSTML  
 LLWTFLLQRRFVKKRGKTGLETAVLLTRINNFME

**Important features of the protein:****Transmembrane domains:**

amino acids 38-60, 83-107, 122-138, 156-173, 189-210, 484-506

**N-glycosylation sites.**

amino acids 349-353, 362-366, 415-419, 442-446

**N-myristoylation sites.**

amino acids 163-169, 413-419, 523-529

**Leucine zipper pattern.**

amino acids 93-115, 109-131

**Glutamine amidotransferases class-II active site.**

amino acids 1-13



[illegible]

**FIGURE 142**

MYLVAGDRGLAGCGHLLVSLGLLLLLLARSGLTRALVCLPCDESKCEEPRNCPGSIVQGVCGCC  
 YTCASQRNESCGLTFTGIYGTCDRGLRCVIRPPLNGDSLTEYEAGVCEDENWTDQQLLGFKPCN  
 ENLIAGCNIINGKCECNTIRTCSNPFEPSPQDMCLSAKRIEEEEKPDCSKARCEVQFSPRCPE  
 DSVLIEGYAPPGECCPLPSRCVCNPAGCLRKVCQPGNLNILVSKASGKPGECDDLIECKPVFG  
 VDCRTVECPPVQQTACPPDSYETQVRLTADGCCTLPTRCECLSGLCGFPVCEVGSTPRIVSRG  
 DGTPGKCCDVFECVNDTKPACVFNNVEYYDGDMMFRMDNCRFCRCQGGVAICFTAQCGEINCER  
 YYVPEGECCPVCEDPVYPFNNPAGCYANGLILAHGDRWREDDCTFCQCVNGERHCVATVCGQT  
 CTNPVKVPGECCPVCEEPTIITVDPACGELSNCCTLTGKDCINGFKRDHNGCRTCQCINTEEL  
 CSERKQGCTLNCPFGFLTDAQNCEICECRPRPKKCRPIICDKYCPLGLLKNKHGCDICRCKKC  
 PELSCSKICPLGFQQDSHGCLICKCREASASAGPPILSGTCLTVDGHGHHKNEESWHDGCRECY  
 CLNGREMCALITCPVPACGNPTIHPGQCCPSCADDFVQKPELSTPSICHAPGGEYFVEGETW  
 NIDSTQCTCHSGRVLCETEVCPPLLCQNPSRTQDSCCPQCTDQPFPSLSRNNVSVNYCKND  
 EGDIFLAAESWKPDVCTSCICIDSVISCFSESCPSVSCERPVLKRGQCCPYCIEDTIPKKVVC  
 HFSGKAYADEERWDLDSCTHCYCLQGQTLCTVSCPPLPCVEPINVEGSCCPMCPEMYVPEPT  
 NIPIEKTNRHGEVDLEVPLWPTPSENDIVHLPRDMGHLQVDYRDNRHLHPSEDSSLDIASVTV  
 PIIICLSIIIIAFLFINQKKQWIPLLCWYRTPTKPSLNNQLVSVDCCKGTRVQVDSSQRMLRI  
 AEPDARFSGFYSMQKQNLQADNFYQTV

**Important features of the protein:****Signal peptide:**

amino acids 1-34

**Transmembrane domain:**

amino acids 940-962

**N-glycosylation sites.**

amino acids 71-75, 113-117, 330-334, 474-478, 746-750

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 992-996

**N-myristoylation site.**amino acids 9-15, 58-64, 61-67, 75-81, 79-85, 362-368, 402-408, 407-413,  
439-445, 492-498, 511-517, 551-557, 558-564, 586-592, 606-612, 625-631,  
845-851**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 52-63, 844-855

**Cell attachment sequence.**

amino acids 314-317

**Leucine zipper pattern.**

amino acids 3-25

**Eukaryotic thiol (cysteine) proteases cysteine active site.**

amino acids 57-69

**VWFC domain proteins.**

amino acids 448-456, 382-390

**C-terminal cystine knot proteins**

amino acids 60-86

**FIGURE 143**

[illegible]

**FIGURE 144**

MVARVGLLLRALQLLLWGHLDQAERGGQELRKEAEAFLEKYGYLNEQVPKAPTSTRFSDAI  
 RAFQWVSQLPVSGVLD RATLRQMTRPRCGVTD TNSYA AWAERISDLFARHRTKMRRKKRFAKQ  
 GNKWKQHL SYRLVNWPEHLPEPAVRGAVRAAFQLWSNVSALEFWEAPATGPADIRLTFFQGD  
 HNDGLGNAFDGPGGALAHAF LPRRGEAHFDQDERWSLSRRRGRNLFVVLAHEIGHTLGLTHSP  
 APRALMAPYYKRLGRDALLSWDDVLAVQSLYGKPLGGSVAVQLPGKLF TDFETWDSYSPQGRR  
 PETQGPKYCHSSFDAITVDRQQQLYIFKGS HFWEVAADGNVSEPRPLQERWVGLPPNIEAAAV  
 SLNDGDFYFFKGGRCWRFRGPKPVWGLPQLCRAGGLPRHPDAALFFPPLRRLILFKGARYYVL  
 ARGGLQVEPYYP RSLQDWGGIPEEVSGALPRPDGSIIFFRDDRYWRLDQAKLQATTSGRWATE  
 LPWMGCWHANSGSALF

**Important features of the protein:****Signal peptide:**

amino acids 1-22

**N-glycosylation sites.**

amino acids 164-168, 355-359

**N-myristoylation sites.**

amino acids 92-98, 153-159, 193-199, 202-208, 288-294, 368-374,  
 509-515

**Amidation site.**

amino acids 312-316

**Neutral zinc metalloproteinases, zinc-binding region signature.**

amino acids 237-247

**Matrixins cysteine switch**

amino acids 231-262, 271-284

**Hemopexin domain protein**

amino acids 66-108, 231-262

CCGGCTAGGGCGCCGGAGCCGCACGCAGCCGCGGGGCTCCGAGAGGCGCGCACTGGGGCTGGGACTGCGCGGGC  
CCGCCGCTGCGAGCGCCACTGAGCGGTGCGCAACTTCGAGAGGCACAGCGCCGGAGCCAGGCGAGCGCTCAGAGA  
CCCGGAGCCAGGGGGCGCGCCGGAGCCTCGTTTCGAGAGCCGGCGCCAGGCACCCACCGCGCTCCGAGTGCCAGG  
CGGCCCTCCGCGACGCGTGGCTTCGCTGCCCCACGGAAGGCACGGGCTGGCGCTGCCGGGCGCCGGGGAGGAC  
GGCAGGAGGAGGCGCGCGCGGAGACGCGCGCGGAGACTGGGGCGAGGAGACGCCCTGGGGAGGAGGCG  
GCCCGAACCAGGCCGCGGGAGC**ATG**GGGGCCCCGGAGCGGAGCTCGGGGCGCGGTGCTGCTGGCACTGCTGCTCTG  
CTGGGACCCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCGTCAGCGCC  
AGCAGAGCCGCTGCCCTACTTCTGTCAGGAGCCACAGGACGCTACATTGTGAAGAACAAGCCTGTGGAGCTCCG  
CTGCCGCGCCTTCCCCGCGACACAGATCTACTTCAAGTGCAACGGCGAGTGGGTGAGCCAGAACGACCACGTAC  
ACAGGAAGGCTTGGATGAGGCCACCGGCTGTCGGGTGCGCGAGGTGCAGATCGAGGTGTGCGGCGACAGGTGGA  
GGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCCTGGCTGGAGCTCCGAGGCACCAAGAGTCCGCC  
AGCCTACGTCCGCATCGCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCTGGACCA  
TGAGGTTCTCCTGCAGTGCCGCCCCGCGGAGGGGGTGCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGT  
CATCGACCCACCCAGGACACCAACTTCTGTCTCACCATCGACCACAACCTCATCATCGCCAGGCCCCGCTGTG  
GGACACTGCCAACTATACCTGCGTGGCCAAGAACATCGTGGCCAAACGCCGGAGCACCCTGCCACCGTCATCGT  
CTACGTGAATCGGGCTGGTCCAGCTGGGCGAGAGTGGTCACTTCTCCAACCGCTGTGGCCGAGGCTGGCAGAA  
GCGCACCCGGACCTGCACCAACCCGCTCCACTCAACGGAGGGGCTTCTGCGAGGGCCAGGCATTCCAGAAGAC  
CGCCTGCACCACCATCTGCCAGTGCATGGGCGCTGGACGGAGTGAGAGCAAGTGGTCAAGCTGCAGCACTGAGTG  
TGCCCACTGGCGTAGCCGCGAGTGCATGGCGCCCCACCCAGAACGGAGGCGCTGACTGCAGCGGGACGCTGCT  
CGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACCTCTAAGCGACCCCAACAGCCACCTGCT  
GGAGGCTCAGGGGATGCGGCGCTGTGATCGGGGGCTCGTGGTGGCCATCTTCGTGGTCTGTGGCAATCCTCATGGC  
GGTGGGGTGTTGGTGTACCGCGCTGAACCTGCGGTGACTTCGCACAGACATCACTGACTCATCTGCTGCCCTGAC  
TGGTGGTTTTCCACCCCGTCAACTTTAAGACGCGAAGGCCAGCAACCCGAGCTCTACACCCCTGTGCTGCCCTC  
TGACCTGACAGCCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCAT  
GACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGG  
GCCAGGCTTGGCAGATGGGGCTGACCTGCTGGGGTCTTGCCGCTGGCACATACCCTAGCGATTCGCCCCGGA  
CACCACCTTCTTGACCTGCGCAGCGCCAGCCTCGTTCCAGCAGCTCTTGCGGCTGCCCGGAGACCCAGGGAG  
ACCGCTGACGCGCACCTTTGGCTGCCTGGTGGGAGGCTCAGCATCCCCGCCACAGGGTGACGTTGCTGGTGCC  
CAATGGAGCCATTTCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAGGCAGAAAGTACCCTCCGCT  
TTCAGAAGGGACCCAGACAGTATTGAGCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCCGCCCGCT  
CATCCTCACCATGCCCCACTGTGCCGAAGTCACTGCCCCGTGACTGGATCTTTAGCTCAAGACCCAGGCCACCA  
GGGCCACTGGGAGGAGGTGTGACCCCTGGATGAGGAGACCCCTGAACACACCCCTGCTACTGCCAGCTGGAGCCAG  
GGCCTGTCACATCCTGCTGGACAGCTGGGACCTACGTGTTCAAGGCGAGTCCATTTCCCGCTCAGCAGTCAA  
GCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCTCTGCACCTCCCTGGAGTACAGCTCCGGGTCTACTGCTGGA  
GGACACGCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGAGCCGAA  
ACCGCTAATGTTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTTGGAGGAG  
CAAGCTGCTGGCCAAATACCAGGAGATCCCCCTTCTATCACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCAC  
TTTCACTCCCTGGAGAGGCACGCTTGGCTTCCACAGAGCTCACCTGCAAGATCTGCGTGCGGCAAGTGGAAAGGGGA  
GGGCCAGATATTCCAGCTGCATACCACTCTGGCAGAGACACCTGCTGGCTCCCTGACACTCTGCTCTGCCCT  
TGGCAGCACTGTCACCACCCAGCTGGGACCTTATGCTTCAAGATGCCCTCCCTGACACTCTGCTCTGCTGCTG  
CAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCT  
GAATTACTTTGCCACCAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGG  
GGACCTCAACAGCCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGG  
GGACTGCT**AG**CCCTCTGGGACAGCGGGCTGGCAGGACTGGCAGGAGGCAGGTGCAGGGAGGCCTGGGGCAGCC  
TCCTGATGGGGATGTTTGGCTCTGCTTCTCCCTGATCTCAGCCAGAGTGGCTCTCTCTCTCTCTCCCCAA  
CCCCCAGACCATGACCAGCCTTAGAAAAATCCATGTACTCTGTTGTTAGAGGGGCCAGAGTCTCTCTCCACCCCC  
GCTCTCTCTCTTGGCTGAGATCTCTGTGCAGGAACCAAGATGGGGCTGAAGCCTCTGGAGGCAGTTGGTTGG  
GGGCGGGCAGGCAGGAGGCCTCCCTCCACCCCCCACCCTCAGCCCCGCAACTTCTGGGTTCGCTGGGTTTTAG  
TTCCGTTCTTCGTTTTCTTCTCTCCGTTATTGATTTCTCCTTTCTCCCTAAGCCCCCTTCTGCTTCCACGCCCTTT  
TCCTCTTTGAAGAGTCAAGTACAATTGACAGAAACTGCTTTCTCTGTCCAAAAGCAAAAAGGCAAAAGGAAAGAA  
AGAAAGCTTCAGACCGCTAGTAAGGCTCAAAAGAAAGAAAAACCAAAAACCAAGGAAAAAGAAAAACCCAG  
TTTCTTAGGAAAACGCAACGATTTATTATCCAGATTATTTGGATAAGTCTTTTTTAAAA

**FIGURE 146**

MGARSGARGALLLALLLCWDPRLSQAGTDSGSEVLPSDFPSAPAEPLPYFLQEPQDAYIVKNK  
PVELRCRAFPATQIYFKCNGEWVSQNDHVTQEGLEATGLRVREVQIEVSRQQVEELFGLEDY  
WCQCVAWSSAGTTKSRRAYVRIAYLRKNFDQEPLGKEVPLDHEVLLQCRPPEGVPVAEVEWLK  
NEDVIDPTQDTNELLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW  
AEWSPCSNRCGRGWQKRTRTCTNPAPLNGGAFCEGQAFQKTACTTICPVDGAWTEWSKWSACS  
TECAHWSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCMQNKKTLSDPNSHLLEASGDAALYA  
GLVVAIFVVVAILMAVGVVVYRRNCRDFTDITDSSAALTGGFHPVNFKTARPSNPQLLHPSV  
PPDLTASAGIYRGPVYALQDSTDKIPMTNSPLLDPLPSLKVKVYSSSTTGSGPGLADGADLLG  
VLPPGTYPSPDFARDTHFLHLRSASLGSQQLLGLPRDPGSSVSGTFGCLGGRLSIPGTGVSLLV  
PNGAIPQGKFYEMYLLINKAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTMPHCAEVSA  
RDWIFQLKTQAHQGHWEVVTLDEETLNTPCYCQLEPRACHILLDQLGTYVFTGESYSRSAVK  
RLQLAVFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVVEPKPLMFKDSYHNRL  
SLHDLPHAHWSKLLAKYQEI PFYHIWSGSQKALHCTFTLERHSLASTEITCKICVRQVEGEG  
QIFQLHTTLAETPAGSLDTLCSAPGSTVTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRM  
LAQKLSMDRYLNYFATKASPTGVILDLWEALQQDDGDLNSLASALEEMGKSEMLVAVATDGDC

**Important features of the protein:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 374-395

**N-glycosylation sites.**

amino acids 222-225, 347-350

**Glycosaminoglycan attachment site.**

amino acids 492-495

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 233-236, 234-237

**Casein kinase II phosphorylation sites.**

amino acids 30-33, 87-90, 251-254, 341-344, 359-362, 629-632, 651-654, 706-709, 757-760, 827-830, 925-928, 941-944

**Tyrosine kinase phosphorylation sites.**

amino acids 216-223, 773-780

**N-myristoylation sites.**

amino acids 2-7, 6-11, 27-32, 96-101, 137-142, 179-184, 247-252, 281-286, 334-339, 379-384, 491-496, 495-500, 509-514, 542-547, 547-552, 550-555, 553-558, 560-565, 611-616, 785-790, 834-839, 844-849

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 541-551

**ATP/GTP-binding site motif A (P-loop).**

amino acids 926-933

**Growth factor and cytokines receptors family signature 2.**

amino acids 306-312

GAGAGGGACAGAGGCTGGAGAAGGATGTATGGCCTGCCCTGGGCTTGCTGTTCCTCCTGAGCCTGAGCCCTT  
ACCTTCCTGACCCC**ATGA**AGCACACACTGGCTCTGCTGGCTCCCCTGCTGGGCCTGGGCCTGGGGCTGGCCCTGA  
GTCAGCTGGCTGCAGGGGCCACAGACTGCAAGTTCCTTGGCCCCGGCAGAGCACCTGACATTACCCCCAGCAGCCA  
GGGCCCCGGTGGCTGGCCCCCTCGAGTTCTGTCGCCAGGACTCCTGGACTCCCTCTATGGCACCGTGCGCCGCTTCC  
TGTCGGTGGTGACGCTCAATCCTTCCCTTCAGAGTTGGTAAGGCCCTACTGAATGAGCTGGCCTCCGTGAAGG  
TGAATGAGGTGGTGCGGTACGAGGGCGGCTACGCTGGTATGCGCTGTGATCGCGGGCCTCTACCTGCTGTGGTGC  
CCACTGCCGGGCTTTGCTTCTGCTGCTGCCGCTGCCACCGGCGCTGCGGGGGACGAGTGAAGACAGAGCACAAAGG  
CGCTGGCCTGTGAGCGCGCGGCCCTCATGGTCTTCTGCTGCTGACCACCCTCTTGCTGCTGATTGGTGTGGTCT  
GTGCCTTTGTACCAACACAGCGCACGCATGAACAGATGGGCCCCAGCATCGAGGCCATGCCTGAGACCCTGCTC  
GCCTCTGGGGCTTGGTCTCTGATGTCCCCCAAGAGCTGCAGGCCGTGGCACAGCAATTCCTCCTGCCCCAGGAGC  
AAGTCTCAGAGGAGCTGGATGGTGTGGTGTGAGCATTGGGAGCGCGATCCACACTCAGCTCAGGAGCTCCGTGT  
ACCCCTTGCTGGCGGCCGTGGGAGTTTGGGCCAGGTCCTGCAAGTCTCCGTGCACCACCTGCAAAACCTTGAATG  
CTACAGTGGTAGAGCTGCAGGCCGGGCAGCAGGACCTGGAGCCAGCCATCCGGGAACACCGGGACCGCCTCCTTG  
AGCTGCTGCAGGAGGCCAGGTGCCAGGGAGATTGTGCAGGGGCCCTGAGCTGGGCCCCGACCCTGGAGCTGGGTG  
CTGACTTTCAGCCAGGTGCCCTCTGTGGACCATGTCTGCAACAGCTAAAAGGTGTCCCCGAGGCCAACTTCTCCA  
GCATGGTCCAGGAGGAGAACAGACCTTCAACGCCCTTCCAGCCCTGGCTGCCATGCAGACATCCAGCGTGGTGC  
AAGAGCTGAAGAAGGCATGGCCCCAGCGCCGGAAGGGGTGAGGACACTGGCTGAAGGGTCCCCGGCTGGAGG  
CAGCTTCCCGCTGGGCCCAGGCATGCAGGAGGTGGAGGAGAGCAGCCGCCCTACTCTCAGGAGGTGCAGAGAT  
ACGAGACCTACAGGTGGATCGTGGGCTGCGTGTCTGCTCCGTGGTCTTATTCGTGGTGTCTTGCAACCTGCTGG  
GCCTCAATCTGGGCATCTGGGGCCTGTCTGCCAGGGACGACCCAGCCACCCAGAAGCCAAGGGCGAGGCTGGAG  
CCCCGTTCCTCTGATGGCAGGTGTGGGCCCTCAGCTTCCCTCTTGCTGCACCCCTCATCCTCCTGGTGTTCGCCACCT  
TCCTGGTGGGTGGCAACGTGCAGCGCTGGTGTGCCGAGCTGGGGAACCGCGAGCTCTTTGAGTTTGACAGACA  
CCCCAGGGAACCTGCCCCCGTCCATGAACCTGTGCAACTCTTGCCCTGAGGAAGAATCATGACATCCACCAAG  
CCTATCAGCAGTGAAGGAAGGGGCAGCGCTCTGGACAGTCTGCAAGCTCAACGACTCCTACGACCTGGAGGAG  
ACCTGGATATCAACCAGTATACCAACAAGCTACGGCAGGAGTTGCAGAGCCTGAAAGTAGACACACAGAGCCTGG  
ACCTGCTGAGCTCAGCCGCCCGCCGCCGACCTGGAGGCCCTGCAGAGCAGTGGGCTTCAGCGCATCCACTACCCCG  
ACTTCTCGTTCAGATCCAGAGGCCCGTGGTGAAGACCAGCATGGAGCAGCTGGCCCCAGGAGTCCAAGGACTGG  
CCCAGGCCCAAGACAATTCGTGTGGGGCAGCGCTGCAGGAGGAGGCCCAAGGACTCAGAAACCTTACCAGG  
AGAAGGTGTCCTCCCCAGCAGAGCCTTGTGGCAAAGCTCAACCTCAGCGTTCAGGGCCCTGGAGTCTCTGCCCGA  
ATCTCCAGCTGGAGACCTCAGATGTCTTAGCCAATGTACCTACCTGAAAGGAGAGCTGCCTGCCTGGGCAGCCA  
GGATGCTGAGGAATGTGAGTGAGTGTCTTCCCTGGCCCGGAGATGGGCTACTTCTCCAGTACGTGGCCTGGGTGA  
GAGAGGAGGTGACTCAGCGCATTCGCACCTGCCAGCCCTTCCGGAGCCCTGGACAACAGCCGTGTGATCCTGT  
GTGACATGATGGCTGACCCCTGGAATGCCCTTCTGGTTCTGCCTGGCATGGTGACCTTCTCTGATCCCCAGCA  
TCATCTTTGCCGTCAAGACCTCCAAATACTTCCGTCTATCCGAAACGCCTCAGCTCCACCAGCTCTGAGGAGA  
CTCAGCTCTTCCACATCCCCCGGGTTACCTCCCTGAAGCTG**TAG**GGCCCTTGTTGGGGTGAGGTGACCCTGAGGCTG  
CCTGTCTTCCCTTTGATTTAGCCTGGGCCACAGGACTTCGGTAGCTCTTGCCCCAGAGCCAGGCTGGCATCCA  
GGCCTGGACATGTCCCCAGTTCCGGCTTACCTGGCCCCACCTTGCCCTGCTCCTTTCCACCCCTTTTCTGCTCAGC  
CCCCATTACCTCAGCTCAGAATCATATGGACTTCTGTCAGCTGCAGAGCCAGCAAGTCCCTACAGGTGTCAAC  
CGTTACCCCCATGCTGGTGGCATCTCACAAGGAAGAGCTGTTCTCCACCTGCTGGAGCCTGGACCCTGGGGTGG  
GACAGAGGCCTCGTCCAACCCCACTCCCCCTTCCCGTGTGTCTTCCCCCTGCCAAGCCTCCCCCTGCCAAGCCTCC  
CCCTGCCCCCTCTCTGAGCCCTCGCCCCCACACCGTCTCTATCTGGCTCCCCCTGGCCCCACTTCCCTCTT  
ATGCCCTTCTTGCCCTTTGCTTCTCCTTCTAGTCCCCCTCTTACCATACTCTCCACTGCTACCTTGCTGGCCCCA  
GAGACACCCCTGCCCAACCAACCACTCAGGTAACGCCACTAATCAGGCAGGGGCCACCATTGGCCTAGGTCTGGG  
CTGGCTGAGGCCCTCGCTCATGGCTCTGAGCCCTCCACTGCCCCAGGGCCTTGGGCCCTGCAGATCTCATC  
CAGGATTTATTGTTGTCCAGTGGGGTGAGGGAGGCCCTGTCTGAAGGCCGAGCCTCCCTGCCGTGACCCCAAGTTAG  
AAATGGGGGTACCAGCACTTAGCTTCTCTGAGTGCTGGCTCCCAAGGAAGGGACCTGGGACCTGGGCCACAGT  
GGGGGCTTGCCCTTACTCTTCAAGGAAGCATCTTCCACAGCCCCCACCACACTTCTTAGGAGTGATCTGGT  
GGCCAGAACAGGATTTTGACACGGCCCCCTTTTATCTCTGCGCATGTGGCCTAGGGCTCATCCCCAGCCCATCCCTGTG  
TCAGCCCTGAGTGTGCGGACCTCGTTCAGAAATGAGGAAGGAGAGAGAAGAGATGGACAGACCTCAGATCC  
ATTAAAGTGTTCTCACTTCAA

**FIGURE 148**

MKHTLALLAPLLGLGLGLALSQLAAGATDCKFLGPAEHLTFTPAARARWLAPRVRAPGLL  
 DSYGTVRRFLSVVQLNPFPSSELVKALLNELASVKVNEVVRYEAGYVVCVAVIAGLYLLL  
 PTAGLCFCCCRCHRRRCGGRVKTEHKALACERAALMVFLLLTLLLLIGVVCAFVTNQRTH  
 EQMGPSIEAMPETLLSLWGLVSDVPQELQAVAQQFSLPQEQVSEELDGVGVSIGSAIHTQ  
 LRSSVYPLLAAGVSLGQVLQVSVHHLQTLNATVVVELQAGQQDLEPAIREHRDRLELLQE  
 ARCQGDCAGALSWARTLELGADFSQVPSVDHVLHQLKGVPEANFSSMVQEENSTFNALPA  
 LAAMQTSSVVQELKKAVAQQPEGVRTLAEGFPGLEAASRWAQALQVEEESRPYLQEVQR  
 YETYRWIVGCVLCSVVLFFVLCNLLGLNLGIWGLSARDDPSHPPEAKGEAGARTLMAGVGL  
 SFLFAAPLILLVFATFLVGGNVQTLVCRSWENGELFEFADTPGNLPPSMNLSQLLGLRKN  
 ISIHQAYQQCKEGAALWTVLQLNDSYDLEEHLDINQYTNKLRQELQSLKVDTSQSLDLLSS  
 AARRDLEALQSSGLQRIHYPDFLVQIQRPVVKTSMEQLAQELQGLAQADNSVLGQRLQE  
 EAQGLRNLHQEKVVPQQSLVAKNLNSVRALESSAPNLQLETSDVLANVTYLGKELPAWAA  
 RILRNVSECFLAREMGYFSQYVAWVREEVTQRIATCQPLSGALDNSRVILCDMMADPWNA  
 FWFCLAWCTFFLIPSIIFAVKTSKYFRPIRKRLSSTSSEETQLFHIPRVTSCLK

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 105-125, 153-173, 428-449, 476-500, 778-797

**N-glycosylation sites:**

amino acids 270-273, 343-347, 352-356, 530-534, 540-546, 563-567,  
 684-688, 707-711, 725-729

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 811-815

**Tyrosine kinase phosphorylation site.**

amino acids 95-103

**N-myristoylation sites.**

amino acids 13-19, 15-21, 17-23, 26-32, 58-64, 124-130, 168-174,  
 228-234, 230-236, 320-326, 338-344, 393-399, 429-435, 446-452,  
 477-483, 500-506, 536-542, 644-650, 761-767

**Phospholipase A2 histidine active site.**

amino acids 129-137

**4Fe-4S ferredoxins, iron-sulfur binding region signature.**

amino acids 126-138

**Mitochondrial energy transfer proteins signature.**

amino acids 80-89



**FIGURE 149**

CACAGCTCCCTTCCCAGGACGTGAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCA  
GCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTG  
CCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAA  
CAAACCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCC  
TTTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCG  
GAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCAC  
AGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAG  
AACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACC  
CCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCAT  
GACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACC  
CAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTTCAGCCTCCTGGCA  
TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGC  
TGGAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCA

## **FIGURE 150**

MRLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVSPNSTNLKGHHVRLCKPC  
KLEPEPRLWVVP GALPQV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-21

**N-glycosylation site.**

amino acids 48-52

**Amidation sites.**

amino acids 23-27, 33-37

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 151**

CACCGGAGGGGCACGCAGCTGACGGAGCTGCGCTGCGTTTCGCCTCGTTTGCCTCGCGCCCTCCA  
 CTGGAGCTGTTTCGCGCCTCCCGGCTCCCACCGCAGCCCACCCGGCAGAGGAGTCGCTACCAGC  
 GCCCAGTGCCTCTGTCTAGTCCGCAAACCTCTTTCGCGCCCGCCCGGGCTGGGCACCAAATAC  
 CAGGCTACCA**ATG**GTCTACAAGACTCTCTTCGCTCTTTGCATCTTAACTGCAGGATGGAGGGTA  
 CAGAGTCTGCCTACATCAGCTCCTTTGTCTGTTTCTCTTCCGACAAACATTGTACCACCGACC  
 ACCATCTGGACTAGCTCTCCACAAACACTGATGCAGACACTGCCTCCCCATCCAACGGCACT  
 CACAACAACCTCGGTGCTCCCAGTTACAGCATCAGCCCCAACATCTCTGCTTCCTAAGAACATT  
 TCCATAGAGTCCAGAGAAGAGGAGATCACCAGCCCAGGTTTGAATTGGGAAGGCACAAACACA  
 GACCCCTCACCTTCTGGGTTCTCGTCAACAAGCGGTGGAGTCCACTTAACAACCACGTTGGAG  
 GAACACAGCTCGGGCACTCCTGAAGCAGGCGTGGCAGCTACACTGTCGCAGTCCGCTGCTGAG  
 CCTCCACACTCATCTCCCCTCAAGCTCCAGCCTCATCACCTCATCCCTATCAACCTCACCA  
 CCTGAGGTCTTTTCTGCCTCCGTTACTACCAACCATAGCTCCACTGTGACCAGCACCCCAACCC  
 ACTGGAGCTCCAACCTGCACCAGAGTCCCCGACAGAGGAGTCCAGCTCTGACCACACACCCACT  
 TCACATGCCACAGCTGAGCCAGTGCCCCAGGAGAAAAACCCCCAACAACTGTGTGAGGCAAA  
 GTGATGTGTGAGCTCATAGACATGGAGACCACCACCACTTTCCCAGGGTGATCATGCAGGAA  
 GTAGAACATGCATTAAAGTTCAGGCAGCATCGCCGCCATTACCGTGACAGTCATTGCCGTGGTG  
 CTGCTGGTGTGTTGGAGTTGCAGCCTACCTAAAAATCAGGCATTCCCTCCTATGGAAGACTTTTG  
 GACGACCATGACTACGGGTCTGGGGAAACTACAACAACCCTCTGTACGATGACTCC**TAA**CAA  
 TGGAATATGGCCTGGGATGAGGATTAAGTGTCTTTATTTATAAGTGCTTATCCAGTAGAATT  
 AATAAGTACCTGATGCGCATTGAACGACAATCTTAAGCCCTGTTTTGTTGGTATGGTTGTTTT  
 TGTTTTCTCTCCCTCTCTCTGGCTGCTACAACTTCCCCCTTTCTGGTACAAGAAGAACCATTCT  
 TTAAAGGTGAGTGGAGGCTGATTTGCAGCTGAAGTGGGCCAGCCTTGCAACAGCCAGGCCAGA  
 CCACCATGGTGAAGGCTTCTTTCCCCACTGCAGGACCCACTTTGAGAAGGATCGAGGAGGAGG  
 ATTTGGGTTGTTTTGTTAGGGGTTACTTTCAGGGGAACATTTTCAATTTGTGTTATTTCTTAAAC  
 TTCTATTTAGGAAATTACATTAAGTATTAATGAGGGGAAAGGAAATGAGCTCTACGAGGATTT  
 CACCTTGCATGGGAGAGAGCAGGGTTTTCTCAGATTCCTTTTTAATCTCTATTTATCTGGTTG  
 TTTCTGACAGGATGCTGCCTGCTTGGCTCTACGAGCTGGAAAGCAGCTTCTTAGCTGCCTAAT  
 TAATGAAAGATGAAAATAGGAAGTGCCCTGGAGGGGGCCAGCAGGTACGGGGCAGAATCTCT  
 CAGGTTGCTGTGGGATCTCAGTGTGCCCTACCTGTTCTCCCCCTCCAGGCCACCTGTCTCTGT  
 AAAGGATGTCTGCTCTGTTCAAAAGGCAGCTGGGATCCCAGCCCACAAGTGATCAGCAGAGTT  
 GCATTTCCAAAGAAAAAGGCTATGAGATGAGCTGAGTTATAGAGAGAAAGGGAGAGGCATGTA  
 CGGTGTGGGGAAGTGGAAGAGAAGCTGGCGGGGGAGAAGGAGGCTAACCTGCACTGAGTACTT  
 CATTAGGACAAGTGAGAATCAGCTATTGATAATGGCCAGAGATATCCACAGCTTGGAGGAGCC  
 CAGAGACTGTTTGCTTTATACCCACACAGCAACTGGTCCACTGCTTTACTGTCTGTTGGATAA  
 TGGCTGTAAATGTTTAAAAAC

## **FIGURE 152**

MVYKTLFALCILTAGWRVQSLPTSAPLSVSLPTNIVPPTTIWTSSPQNTDADTASPSNGTHNN  
SVLPVTASAPTSLLPKNISIESREEEITSPGSNWEGTNTDPSPSGFSSTSGGVHLTTTLEEHS  
SGTPEAGVAATLSQSAAEPPTLISPQAPASSPSSLSTSPPEVFSASVTTNHSSTVTSTQPTGA  
PTAPESPTEESSDHTPTSHATAEPVPQEKTPPTTVSGKVMCELIDMETTTTFPRVIMQEVEH  
ALSSGSIAAITVTVIAVLLVFGVAAYLKIRHSSYGRLLDDHDYGSWGNYNPLYDDS

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-20

#### **Transmembrane domain:**

amino acids 258-278

#### **N-glycosylation sites.**

amino acids 58-61, 62-65, 80-83, 176-179

#### **Casein kinase II phosphorylation sites.**

amino acids 49-52, 85-88, 95-98, 100-103, 120-123, 121-124, 141-144, 164-167, 191-194, 195-198, 200-203

#### **Tyrosine kinase phosphorylation site.**

amino acids 289-296

#### **N-myristoylation sites.**

amino acids 59-64, 115-120, 128-133, 133-138, 257-262, 297-302

**FIGURE 153**

ACGTCACTGTCTTGAAGCAGCAGTAGCCTGGGAAGTGAGGCAGGAGGAATTGAGAGGCAGGAA  
GGGNGCTGGAGACACAGCTGAGCCTGGAAATGAGAGTGGGCATCGCCGTGGTCATCATGACTC  
CTCTGCGGCGTGGTCACCATGTTGGTTCACTGTGTTGGGCTCTTATTGACGGGTCTCCTGCTA  
GGCCTGACCTTGGGTGCCGGAGCCCTGCTGGCTTCTGAGCCTATCTACCAACCACCTTCAGCC  
TGGGTGCCAGCTGGGGGGCTGGTGGGGCTGGCGCTGCTGGGAGCCCTGCTCACACTTCGGTGG  
CCACGTCCATTACAGTTCTGGGCACAACCCTGCTGGGTTCTGCAGTGCTTGTGGCCTGTGTT  
GACTACTTCCTGGAGGGGGCTGGCACTGGGGAGTTGGCTGGGCCAACGCCTGCAGACACTTCCA  
GCCTTGCCCTTCTCTCTGCTTGATATAGCTGGGTCTTACTGGGGATCTGGCCAGCCTTGGGGGCC  
CTTGGAGCCCTGGCCCAGTGGAAGCTCGTGCCTGAGGAACATGGAGGCCACGCTAATGGGTCT  
GTTCTTGGTTTCCCAGATGCATAAAGGAAGACATATCCCTCCCCTGGGCAGCAAGGCTACAAT  
GGGAGGGAGGGAGAACATGGGAGCATGTGAATAAAATGGCATTAATACTGAAAAAAAAAAAAA  
AAA

## **FIGURE 154**

MLVHCVGLLLTGLLLGLTLGAGALLASEPIYQPPSAWVPAGGLVGLALLGALLTLRWPRPFTV  
LGTTLGSAVLVACVDYFLEGLALGSWLGQRLQTLPALPSLC

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 38-55, 60-78

**N-myristoylation sites.**

amino acids 7-13, 12-18, 16-22, 22-28, 41-47, 50-56, 84-90, 88-94

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 67-78

**FIGURE 155**

TTGCAATTAAAGGAGTCGGGTCTCTAACTGTTTGATCTGTTTTTTTTCCCTTCTGAGCAATGGAGC  
TTACCATCTTTATCCTGAGACTGGCCATTTACATCCTGACATTTCCCTTGTACCTGCTGAACT  
TTCTGGGCTTGTGGAGCTGGATATGCAAAAAATGGTTCCCCTACTTCTTGGTGAGGTTCACTG  
TGATATACAACGAACAGATGGCAAGCAAGAAGCGGGAGCTCTTCAGTAACCTGCAGGAGTTTG  
CGGGCCCCCTCCGGGAAACTCTCCCTGCTGGAAGTGGGCTGTGGCACGGGGGCCAACTTCAAGT  
TCTACCCACCTGGGTGCAGGGTGACCTGTATTGACCCCAACCCAACTTTGAGAAGTTTTTTGA  
TCAAGAGCATTGCAGAGAACCGACACCTGCAGTTTGAGCGCTTTGTGGTAGCTGCCGGGGAGA  
ACATGCACCAGGTGGCTGATGGCTCTGTGGATGTGGTGGTCTGCACCCTGGTGCTGTGCTCTG  
TGAAGAACCAGGAGCGGATTCTCCGCGAGGTGTGCAGAGTGCTGAGACCGGGAGGGGCTTTCT  
ATTTTCATGGAGCATGTGGCAGCTGAGTGTTCGACTTGGAATTACTTCTGGCAACAAGTCTTG  
ATCCTGCCTGGCACCTTCTGTTTGATGGGTGCAACCTGACCAGAGAGAGCTGGAAGGCCCTGG  
AGCGGGCCAGCTTCTCTAAGCTGAAGCTGCAGCACATCCAGGCCCCACTGTCCTGGGAGTTGG  
TGCGCCCTCATATCTATGGATATGCTGTGAAATAGTGTGAGCTGGCAGTTAAGAGCTGAATGG  
CTCAAAGAATTTAAAGCTTCAGTTTTACATTTAAAATGCTAAGTGGGAGAAGAGAAACCTTTT  
TTTTTGGGGGGCGGTTTTTTTTTGGTTTTGTTGTTGGTTTTTTTTTTTTTTTTTTGGCAGGAGAATCTC  
TTGAACCCAGAAGGCGAAGGTTGCAGTGAACCGAGATCATGCCATTGTACTCTAGCCTGGGTG  
ACAAGAGCAAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGAAGTAGAGACAGGGAGAC  
GGGGTCTCACTGTGTTGCCTAGGCCGGTCTTGAACCTCCTGGGCTCAAGTGATTCTCCACCTT  
GACCTCCTAAATTGTTGGGATTACAGGTGTGAGACAGTGCACCTGGCCGAAATAGCTCAAGTT  
TCTGAAAAACAAATCTGAATCTATTTGTTATTCTTAGCGTCACTGGTCTGGCTTTCAGAATTA  
ACATACAAGGTTGCCACACCTAGTTCTGCCAGCTTTATGTCTTTTATTCCAGTATTCCACCA  
AAGTTTGTTTTCTGCAATCCAGTTCTCAAGTCTTAAGATAAAGATTGTACTTGACAGTTTAG  
TATATCCATAAAACTATTTGAGGTGGTTAAGGTTCTTGGGTTCATTTTCTTAATACTTTGCT  
GAATATTGTAGATTGTAGGCAATGAAAAAGTCTACTAAATTAGGAAAACCTTGAATAATTAGG  
TATCCTAGGTAAGAGCCCCTAAACATCAAGCAATCTGTGAGTCTGTAAAGAAATAAATATTTT  
TTGGATTATTCTTATCTAATTCCACCCCTGTTGGAAGATGATTTCTTTGTTCTTTGCAACTAT  
GGAAGCTGTGAAAATCATCACAAGTGCCTCTGAAAGCGAGTGTAGGTTGGTTAGAGGGTTTA  
ATATTTTCTGCAATGGTTTGTAGGAATTTTAATAAATGTAGTATATTTTCTGAGATGATTTTG  
TAAAGTACTATTTTAAATATCAAATCAACCAATAAATTCACATTTGTGTTAGGAACAAAA

## FIGURE 156

MELTIFILRLAIYILTFPLYLLNFLGLWSWICKKWFYFLVRFTVIYNEQMASKKRELFSLNQ  
EFAGPSGKLSLLEVCGGTGANFKFYPPGCRVTCIDPNPNFEKFLIKSIAENRHLQFERFVVA  
GENMHQVADGSVDVVVCTLVLCVKNQERILREVCRLVRPGGAFYFMEHVAAECSTWNYFWQQ  
VLDPAWHLLFDGCNLTRESWKALERASFSLKLQHIQAPLSWELVRPHIYGYAVK

Signal peptide:

amino acids 1-29

N-glycosylation site.

amino acids 203-207

N-myristoylation sites.

amino acids 78-84, 80-86, 91-97, 201-207



**FIGURE 157**

CCGCTGAGATGTACGAACTTCCGGTTCTCCGGGCAGCTGCCACTGCTGTAGCTTCTGCCACCT  
GCCACGACCGGGCCTCTCCCTGGCGTTTGGTCACCTCTGCTTCATTCTCCACCGCGCCTATGG  
TCCCTCTTGGAGCCAGCGTGGCGGGCCTGGCGGGCTCCCGGGTGGTGAGAGAGCGGTCCGGGAA  
CG**ATGA**AAGGCCTCGCAGTGCTGCTGCTGTCTCAGCCACCTCTTGGCTTCCGTCCTCCTCCTGC  
TGTTGCTGCCTGAACTAAGCGGGCCCCTGGCAGTCCTGCTGCAGGCAGCCGAGGCCGCGCCAG  
GTCTTGGGCCTCCTGACCCTAGACCACGGACATTACCGCCGCTGCCACCGGGCCCTACCCCTG  
CCCAGCAGCCGGGCGGTGGTCTGGCTGAAGCTGCGGGGCCGCGGGGCTCCGAGGGAGGCAATG  
GCAGCAACCCCTGTGGCCGGGCTTGAGACGGACGATCACGGAGGGAAGGCCGGGGAAGGCTCGG  
TGGGTGGCGGCCTTGCTGTGAGCCCCAACCCCTGGCGACAAGCCCATGACCCAGCGGGCCCTGA  
CCGTGTTGATGGTGGTGAGCGGCGCGGTGCTGGTGTACTTCGTGGTCAGGACGGTCAGGATGA  
GAAGAAGAAACCGAAAGACTAGGAGATATGGAGTTTTTGGACACTAACATAGAAAATATGGAAT  
TGACACCTTTAGAACAGGATGATGAGGATGATGACAACACGTTGTTTGATGCCAATCATCCTC  
GAAGA**TAA**GAAATGTGCCTTTTGATGAAAGAACTTTATCTTTCTACAATGAAGAGTGGAATTTT  
TATGTTTAAGGAATAAGAAGCCACTATATCAATGTTGGGGGGTATTTAAGTTACATATATTT  
TAACAACCTTTAATTTGCTGTTGCAATAAATACCGTATCCTTTTATTATATCTTTATATGTAT  
AGAAGTACTCTATTAATGGGCTCAGAGATGTTGGGGATAAAGTATACTGTAATAATTTATCTG  
TTTGAAAATTACTATAAAACGGTGTTTTCTGGTCGGTTTTTGTTCCTGCTTACCATATGATT  
GTAAATTGTTTTATGTATTAATCAGTTAATGCTAATTATTTTTGCTGATGTCATATGTTAAAG  
AGCTATAAATTCCAACAACCACTGGTGTGTAAAAATAATTTAAATTTTCCTTTACTGAAAGG  
TATTTCCCATTTTTGTGGGGAAAAGAAGCCAAATTTATTACTTTGTGTTGGGGTTTTTAAAT  
ATTAAGAAATGTCTAAGTTATTGTTTGCAAAACAATAAATATGATTTTAAATTCTCTTAAAA  
AAAA

## **FIGURE 158**

MKASQCCCCLSHLLASVLLLLLLPELSGPLAVLLQAAEAAPGLGPPDPRPRTLPPPLPPGPTPA  
QQPGRGLAEAAGPRGSEGGNGSNPVAGLETDDHGGKAGEGSVGGGLAVSPNPGDKPMTQRALT  
VLMVVSGAVLVYFVVRTVRMRRNRKTRRYGVLDTNIE NMELTPLEQDDEDDNTLFDANHPRR

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 124-140

**N-glycosylation site.**

amino acids 83-87

**N-myristoylation sites.**

amino acids 69-75, 78-84, 81-87, 97-103, 103-109, 106-112,  
157-160

GCTGCGAGCGCGGACGGCGCTACACCA**ATGG**GGCCGGCTGCTGCGGGCCGCGGCTGCCCGCGCTG  
 CTTTCGCCGCTGCTGCTTCTGCTGGTTGGGGGAGCGTTCCCTGGGTGCCTGTGTGGCTGGGTCT  
 GATGAGCCTGGCCAGAGGGCCTCACCTCCACCTCCCTGCTAGACCTCCTGCTGCCACTGGC  
 TTGGAGCCACTGGACTCAGAGGAGCCTAGTGAGACCATGGGCCTGGGAGCTGGGCTGGGAGCC  
 TCTGGCTCAGGCTTCCCCAGCGAAGAGAATGAAGAGTCTCGGATTCTGCAGCCACCACAGTAC  
 TTCTGGGAAGAGGAGGAAGAGCTGAATGACTCAAGTCTGGACCTGGGACCCACTGCAGATTAT  
 GTTTTTTCTGACTTAAGTGAAGAAGGCAGGTTCCATTGAAGACACTAGCCAGGCTCAAGAGCTG  
 CCAAACCTCCCCTCTCCCTTGCCCAAGATGAATCTGGTTGAGCCTCCCTGGCATATGCCTCCC  
 AGAGAGGAGGAAGAAGAGGAAGAGGAAGAGGAGGAGGGAGAAGGAAGAGGTAGAGAAACAA  
 GAGGAGGAGGAAGAGGAGGAGCTGCTCCCTGTGAATGGATCCCAAGAAGAAGCCAAGCCTCAG  
 GTCCGTGACTTTTTCTCTCACCAGCAGCAGCCAGACCCAGGGGCCACCAAAAGCAGGCATGAA  
 GACTCCGGGGACCAGGCCTCATCAGGTGTGGAGGTGGAGAGCAGCATGGGGCCCAGCTTGCTG  
 CTGCCTTCAGTCACCCCAACTACAGTGACTCCGGGGGACCAGGACTCCACCAGCCAAGAGGCA  
 GAGGCCACAGTGCTGCCAGCTGCAGGGCTTGGGGTAGAGTTTCGAGGCTCCTCAGGAAGCAAGC  
 GAGGAAGCCACTGCAGGAGCAGCTGGTTTTGTCTGGCCAGCACGAGGAGGTGCCGGCCTTGCTT  
 TCATTCCCTCAAACCACAGCTCCCACTGGGGCCGAGCACCCAGATGAAGATCCCCTTGCTCT  
 AGAACCTCAGCCTCTTCCCCACTGGCCCTGGAGACATGGAAGTACACCTTCTCTGCTACC  
 TTGGGACAAGAAGATCTCAACCAGCAGCTCCTAGAAGGGCAGGCAGCTGAAGCTCAATCCAGG  
 ATACCCTGGGATTCTACGCAGGTGATCTGCAAGGACTGGAGCAATCTGGCTGGGAAAACTAC  
 ATCATTCTGAACATGACAGAGAACATAGACTGTGAGGTGTTCGGCAGCACCCGGGGGCCACAG  
 CTCTTGCCCTGGTGGAAGAGGTGCTGCCCCGCCATGGCAGTGCCACCATGGGGCCTGGCAC  
 ATCTCTCTGAGCAAGCCCAGCGAGAAGGAGCAGCACCTTCTCATGACACTGGTGGGCGAGCAG  
 GGGGTGGTGCCCACTCAAGATGTCCTTTCCATGCTGGGTGACATCCGCAGGAGCCTGGAGGAG  
 ATTGGCATCCAGAACTATTCCACAACCAGCAGCTGCCAGGCGCGGGCCAGCCAGGTGCGCAGC  
 GACTACGGCAGCTCTTCGTGGTGCTGGTGGTCATTGGGGCCATCTGCATCATCATCATTTGCG  
 CTTGGCCTGCTCTACAAGTCTGGCAGCGCCGGCTGCCAAGCTCAAGCACGTGTGCGACGGC  
 GAGGAGCTGCGCTTCGTGGAGAACGGCTGCCACGACAACCCACGCTGGACGTGGCCAGCGAC  
 AGCCAGTCGGAGATGCAGGAGAAGCACCCACGCTGAACGGCGGGCGGGGCCCTCAACGGCCCCG  
 GGGAGCTGGGGGGCGCTCATGGGGGGCAAGCGGGACCCCGAGGACTCGGACGTGTTTCGAGGAG  
 GACACGCACCTG**TGA**GCGCAGCCGAGGCGCAGGCCGAGTGGGCCGCCAGGACCAAGCGAGGTG  
 GACCCCGAAACGGACGGCCCCGAGCCCCGCACCAGCCCCGCGCCTACCCGGGCCGCCCGCGCG  
 CCTGGCCCTCGGCGCGGGGCTCCTTCCCGCTTCCCCGACTTCACACGGCGGGCTTCGGACCAAC  
 TCCCTCACTCCCGCCCCGAGGGGCAGGCCTCAAAGCCCCGCTTGGCCCCGCTTTCGCGCCCTG  
 AACCCCGGCCCCGCGGGCGGGCGGGCGGCGCTTCTGCGCCCCGGGACTCAATTAAACCCGCC  
 GGAGACCACGCCGGGCCAGCAAAA

**FIGURE 160**

MGRLLRAARLPPLLSPLLLLLLVGGAFLGACVAGSDEPGPEGLTSTSLDLLLPTGLEPLDSEE  
PSETMGLGAGLGASGSGFPSEENEESRILQPPQYFWEEEEELNDSSLDLGPTADYVFPDLTEK  
AGSIEDTSQAQELPNLPSPLPKMNLVEPPWHMPPREEEEEEEEEEREKEEVEKQEEEEEEEL  
LPVNGSQEEAKPQVRDFSLTSSSQTPGATKSRHEDSGDQASSGVEVESSMGPSLLLPSVTPTT  
VTPGDQDSTSQEAETVLPAAGLGVEFEAPQEASEEATAGAAGLSGQHEEVPALPSFPQTAP  
SGAEHPDEDPLGSRTSASSPLAPGDMELTPSSATLGQEDLNQQLLEGQAEEAQSRIPWDSTQV  
ICKDWSNLAGKNYIIILNMTENIDCEVFRQHRGPQLLALVEEVLPRHGSGHHGAWHISLSKPSE  
KEQHLLMTLVGEQGVVPTQDVL SMLGDIRRSLEEIGIQNYSTTSSCQARASQVRSYGTFLFVV  
LVVIGAICIIIIALGLLYNCWQRRPKLKHVSHGEELRFVENGCHDNPTLDVASDSQSEMQEK  
HPSLNGGGALNGPGSWGALMGGKRDPESDVFEEETHL

**Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 499-521

**N-glycosylation sites.**

amino acids 106-110, 193-197, 395-399, 480-484

**Glycosaminoglycan attachment site.**

amino acids 77-81

**N-myristoylation sites.**

amino acids 24-30, 28-34, 41-47, 69-75, 71-77, 73-79, 75-81,  
216-222, 327-333, 455-461, 519-525, 574-580, 581-587, 584-590

**Amidation site.**

amino acids 588-592

**FIGURE 161**

CCAGGGCGGAGCGCAGCTGCGCCGGGCTTGGGCGCCTGGGGCCGCGCTCCCCACCGTCGTTT  
TCCCCACCGAGGCCGAGGCGTCCCGGAGTCAATGGCCGGCCTGAACTGCGGGGTCTCTATCGCA  
CTGCTAGGGGTTTCTGCTGCTGGGTGCGGCGCGCCTGCCGCGCGGGGCAGAAGCTTTTGAGATT  
GCTCTGCCACGAGAAAGCAACATTACAGTTCTCATAAAGCTGGGGACCCCGACTCTGCTGGCA  
AAACCCTGTTACATCGTCATTTCTAAAAGACATATAACCATGTTGTCCATCAAGTCTGGAGAA  
AGAATAGTCTTTACCTTTAGCTGCCAGAGTCCTGAGAATCACTTTGTCATAGAGATCCAGAAA  
AATATTGACTGTATGTCAGGCCCATGTCCTTTTGGGGAGGTTGAGCTTCAGCCCTCGACATCG  
TTGTTGCCTACCCTCAACAGAACTTTTCATCTGGGATGTCAAAGCTCATAAGAGCATCGGTTTA  
GAGCTGCAGTTTTTCCATCCCTCGCCTGAGGCAGATCGGTCCGGGTGAGAGCTGCCAGACGGA  
GTCACTCACTCCATCAGCGGCCGAATCGATGCCACCGTGGTCAGGATCGGAACCTTCTGCAGC  
AATGGCACTGTGTCCCGGATCAAGATGCAAGAAGGAGTGAAAATGGCCTTACACCTCCCATGG  
TTCCACCCCAGAAATGTCTCCGGCTTCAGCATTGCAAACCGCTCATCTATAAAACGTCTGTGC  
ATCATCGAGTCTGTGTTTGAGGGTGAAGGCTCAGCAACCCTGATGTCTGCCAACTACCCAGAA  
GGCTTCCCTGAGGATGAGCTCATGACGTGGCAGTTTGTCGTTTCTGCACACCTGCGGGCCAGC  
GTCTCCTTCCTCAACTTCAACCTCTCCAACCTGTGAGAGGAAGGAGGAGCGGGTTGAATACTAC  
ATCCCGGGCTCCACCACCAACCCCGAGGTGTTCAAGCTGGAGGACAAGCAGCCTGGGAACATG  
GCGGGGAACCTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAAGTCCAGGGATCCTC  
CGGCTGCAGTTCGAAGTTTTTGGTCCAACATCCACAAAATGAAAGCAGTGAGTGAGCCCCACTT  
TCCTTTTTCTTCCTCCTCCAGCACCTTCGTTGTTTCCTGGGTAGTCTGCCTGGGTGAGGCTCC  
CTTCCTGTTTCTCATCTGTGGCTTCTGAAACACTTAGACTCTGGACCCAGCAAGAGTTTCAGG  
AAGTGGGTTGCTAGGCAGTTAGACAGGCTTGTTGGTGAACACCCGGTATGTAGTTCATTTCA  
GCACAATAAAAAGAAATCTTGCATTCAAGATGCTAAATTGTTTTTAACGAAAA

**FIGURE 162**

MAGLNCGVSIALLGVLLLGAARLPRGAFAFEIALPRESNITVLIKLGTPTLLAKPCYIVISKR  
HITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCPFGEVQLQPSTSLPLNRTFI  
WDVKAHKSIGLELQFSIPRLRQIGPGESCPDGVTHSISGRIDATVVRIGTFCSNGTVSRIKMQ  
EGVKMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFEGEGSATLMSANYPEGFPEDELMTW  
QFVVPAHLRASVSFLNFNLSNCERKEERVEYYIPGSTTNPEVFKLEDKQPGNMAGNFNLSLQG  
CDQDAQSPGILRLQFQVLVQHPQNESSE

**Signal peptide:**

amino acids 1-29

**N-glycosylation sites.**

amino acids 39-43, 122-126, 180-184, 205-209, 213-217, 270-274,  
310-314, 339-343

**Tyrosine kinase phosphorylation site.**

amino acids 276-284

**N-myristoylation sites.**

amino acids 3-9, 7-13, 158-164, 175-181, 191-197, 303-309

**FIGURE 163**

CAACACACACCTGGGGAATTGCTGGCCTGACTTCTGACCCCTGACTCCTCATACCCCTTCCTC  
CAGAGCATGACATTTGACCACCAACTGAAACCTGACCTCTGACCCAGACCACTGGCCCTTCC  
CCCGCCCTGTGGTGACTTCATAAAGGTTACTAGCTTCTCCCCTGGCCTTGAGACCCACACG**AT**  
**GG**CCCTGCTGGCTCTGGCCAGTGCCGTCCCCTCTGCCCTGCTGGCCCTGGCTGTCTTCAGGGT  
GCCCCCTGGGCCTGTCTCCTCTGCTTCACAACCTACTCTGAGCGCCTCCGCATCTGCCAGAT  
GTTTGTTGGGATGCGGAGCCCCAAGCTTGAAGAGTGTGAGGAGGCCTTCACGGCCGCCTTCCA  
GGGCCTCTCTGACACCGAAATCAGTGAGGAGACCATCCACACTTCATCAGTGTCTTGGGGAAG  
GTGCAGAGGGAGGGCAGGAGAGGCCAGAGGGTCAGGCTGAGGGACAGACAGAGAGAAACAGT  
CAGAGGAGAAAGGCTCAAAGACCATGAGAACAACAGAGACTTAGGGACAGAGAGACACAGACA  
GGGGAAGACAGCAGGGCAAAGACTCAGAGAGGGGGAGGATGGAGAGTCAGAGAGGGGGAAGATGG  
AGACTCAGAGAGAGGGGAGGATGGAGACTCAGAGAGAGAGGAAGATGGAGACTCAGAGGGGAAA  
GATGGAGACTCAGGAGTATGGAGAGTCAGAGAGGGGAGGATGGACACTCAGGGGAGGATGGAG  
AGTCAGGAGGATGGAGACTCATAGAAAGGGGAGGATGGAGAGTCAGGAGAGGTTGGAGACTGG  
AGAGGGAATAGAGACCCAGAAAGGGGAGGATGGAGACTCAGAGGGTGGAAGATGGAGACTCAA  
AGAGGATGGAACCCAGAGAGAGGAGGACAGAGAT**TGA**GGCAGAGACTAGGGGAAGCAGGATAG  
CGACTGGTTCGGGGGCAGAGACTCAGGGAGGATAGAGACTCACAGAGAGGTGAGGATAGAGACT  
TGAGGAGGGACTCAGGAAGCATAGCGACTGTGGGGCAAAGAGTCAGAGAGGGGAGGATACAGAC  
TTGGGAGGGCAGAGACTCAGAAACAGAATGTTTCGCATTAGGGACATGGTGTTCGGGGAGCTG  
CCTCCCCCAGCCCCTGCTCCCTCCCTCACCGCCAGACTATGATGAGAGAAGCCACCTGCATGA  
CACCTTCACCCAGATGACCCATGCCCTGCAGGAGCTGGCTGCTGCCCAGGGATCCTTTGAGGT  
TGCCTTCCCTGATGCTGCAGAGAAAATGAAGAAGGTCATTACACAGCTTAAAGAAGCCCAGGC  
TTGCATCCCTCCCTGCGGTCTCCAGGAGTTCGCCCCGGCGTTTCTCTGCAGCGGGTGCTACTC  
TAGGGTCTGCGACCTCCCGCTGGACTGCCCAGTTCAGGATGTGACAGTGACTCGGGGCGACCA  
GGCTATGTTTTCTTGCATCGTAACTTCCAGCTGCCAAAGGAGGAGATCACCTATTCTTGGA  
GTTTCGCAGGAGGAGGTCTCCGGACTCAGGACTTGTCTTATTTCCGAGATATGCCGCGGGCCGA  
AGGATACCTGGCGCGGATCCGGCCGGCTCAGCTCACGCACCGCGGGACGTTCTCCTGCGTGAT  
CAAGCAAGACCAGCGCCCCCTGGCCCGGCTCTACTTCTTTCTTAACGTCCTCGGGGCCCTCGC  
ATCAGCGAGTGCGACAGTGTTGGCGTGGTGAGTTCTGGGGACTCCGGAGCCCCAGCATCTAGC  
TCCCCGCTGTCTCAGATCCCACCGAGAAGTCTGGGTTCCAGCAACCTCCAACCCAGGAGGAT  
GTTCTTTTCGATGGTACTGCAGTGGCAACTAACAAAGGTATCTTTCTCTCCTTCCCTATCCTATT  
TCCATCCTGAAAATAAAGAATATATTTCAACTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAA

## **FIGURE 164**

MA LLALASAVPSALLALAVFRVPAWACLLCFTTYSERLRICQMFVGMRS PKLEECEEAF TA AF  
QGLSDTEISEETIHTSSVSWGRRCRGRAGEAQRVRLRDRQRET VRGERLKD HENNRDLGTERHR  
QGKTAGQRLREGRMESQRGEDGD SERGEDGD SEREEDGD SEGKMETQEYGESERGGWTLRGGW  
RVRRMETHRKGRMESQERLETGEG IETQKGEDGDSEGGRWRLKEDGNPERGGQR

**Signal peptide:**

amino acids 1-26

**N-myristoylation site.**

amino acids 65-71





**FIGURE 166**

MELSDVTLIEGVGNEVMVVAGVVVLILALVLAWLSTYVADSGSNQLLGAIVSAGDTSVLHLGH  
VDHLVAGQGNPEPTELPHPSEGNDKEAEAGEGRGDSTGEAGAGGGVEPSLEHLLDIQGLPKR  
QAGAGSSSPEAPLRSEDSTCLPPSPGLITVRLKFLNDTEELAVARPEDTVGALKSKYFPGQES  
QMKLIYQGRLQDPARTLRSLNITDNCVIHCHRSPPGSAVPGPSASLAPSATEPPSLGVNVGS  
LMVPVFVVLGLGVVWYFRINYRQFFTAPATVSLVGVTVFFSFLVFGMYGR

Signal peptide:

amino acids 1-36

**Transmembrane domains:**

amino acids 246-267, 275-301

N-glycosylation sites.

amino acids 162-166, 211-215

N-myristoylation sites.

amino acids 48-54, 105-111, 109-115, 129-135, 177-183, 247-253

Cell attachment sequence.

amino acids 97-100



## **FIGURE 168**

MWFMYLLSWLSLFIQVAFITLAVAAGLYYLAELIEEYTVATSRIIKYMIWFSTAVLIGLYVFE  
RFPTSMIGVGLFTNLVYFGLLQTFPFIMLTSPNFILSCGLVVVNHYLAFQFFAEYYPFSEVL  
AYFTFCLWIIIPFAFFVSLSAGENVLPSTMQPGDDVVSNYFTKGKRGKRLGILVVFSFIKEAIL  
PSRQKIY

**Signal peptide:**

amino acids 1-25

**Transmembrane domain:**

amino acids 126-146

**Casein kinase II phosphorylation site.**

amino acids 145-148

**N-myristoylation sites.**

amino acids 73-78, 82-87

**Amidation sites.**

amino acids 168-171, 171-174

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 91-101

**FIGURE 169**

CAAAGCCCTACCCTCACCATTACCCAGGTCCTGTGGGAAGAGCAGCGTGGAGGTGGGCTGAGG  
 TTAGAAGGTGCAGAGCGTGGAAGAAGATTGTGAGCTGAGTATTGGACATCTGTTCTTGAATAG  
 TCCCTGGGCCTGCCATAGGAAAGGAAGTTCTCCAGGGTTACAGTTCTTATCCGCGTGAATACA  
 CATGGCTCTGTTACGAAAAATTAATCAGGTGCTGCTGTTCCCTTCTGATCGTGACCCTCTGTGT  
 GATTCTGTATAAGAAAGTTCATAAGGGGACTGTGCCAAGAATGACGCAGATGATGAATCCGA  
 GACTCCTGAAGAACTGGAAGAAGAGATTCCCTGTGGTGATTTGTGCTGCAGCAGGGAGGATGGG  
 TGCCACTATGGCTGCCATCAATAGCATCTACAGCAACACTGACGCCAACATCTTGTTCTATGT  
 AGTGGGACTCCGGAATACTCTGACTCGAATACGAAAATGGATTGAACATTCCAACTGAGAGA  
 AATAAACTTTAAAATCGTGGAATTCAACCCGATGGTCCTCAAAGGGAAGATCAGACCAGACTC  
 ATCGAGGCCTGAATTGCTCCAGCCTCTGAACTTTGTTGATTTTATCTCCCTCTACTTATCCA  
 CCAACACGAGAAAGTCATCTATTTGGACGATGATGTAATTGTACAAGGTGATATCCAAGAACT  
 GTATGACACCACCTTGCCCTGGGCCACGCGGCGGCTTTCTCAGATGACTGCGATTTGCCCTC  
 TGCTCAGGACATAAACAGACTCGTGGGACTTCAGAACACATATATGGGCTATCTGGACTACCG  
 GAAGAAGGCCATCAAGGACCTTGGCATCAGCCCCAGCACCTGCTCTTTCAATCCTGGTGTGAT  
 TGTTGCCAACATGACAGAAATGGAAGCACCAGCGCATCACCAAGCAATTGGAGAAATGGATGCA  
 AAAGAATGTGGAGGAAAACCTCTATAGCAGCTCCCTGGGAGGAGGGGTGGCCACCTCCCCAAT  
 GCTGATTGTGTTTCATGGGAAATATTCCACAATTAACCCCTGTGGCACATAAGGCACCTGGG  
 CTGGAATCCAGATGCCAGATATTCGGAGCATTTTCTGCAGGAAGCTAAATTAATCCACTGGAA  
 TGGAAGACATAAACCTTGGGACTTCCCTAGTGTTTACAACGACTTATGGGAAAGCTGGTTTGT  
 TCCTGACCCTGCAGGGATATTTAACTCAATCACCATAGCTGATATAACTCTACCCTTAAAT  
 ATTCCTGTATAGAAATGTGGAATTGTCCCTTTGTAGCCAACTATAACATTGTTCTTTATGAA  
 TATTACCTTTGATACATATGATCCACAATATAAAAACCAAAACTACTGTGTGCAAATTATAC  
 CTTGGACCATATAGGCATTGATTAACCTCTTTAAGTACATGTGATAACTATGGAAATCAAGAT  
 TATGTGACTGAAAAACATAAAGGAAGAGACCCATCTAGATAACAGCAATCAACCTGCTTAATT  
 CTGAATGACAATTATATCCACAAATTTTAAAACCTTCTACATGTATTTTTCACATGAAGATCT  
 CCTTAACAGGTTGCCAACCTTTTCTTTTATAAACTATTACATTTAAATATGGACGTCTGAA  
 AAATAAAATATTCATCATTTTTTAAAA

## **FIGURE 170**

MALLRKINQVLLFLLIVTLCVILYKKVHKGTVPKNDADDESETPEEEEEEIPVVICAAAGRMG  
ATMAAINSIYSNTDANILFYVGLRNTLTRIRKWIEHSLREINFKIVEFNPMLKGKIRPDS  
SRPELLQPLNFVRFYLPLLIHQHEKVIYLDLDDVIVQGDIQELYDTTLALGHAAAFSDDCDLPS  
AQDINRLVGLQNTYMGYLDYRKKAIKDLGISPSTCSFNPGVIVANMTEWKHQIRITKQLEKWMQ  
KNVEENLYSSSLGGGVATSPMLIVFHGKYSTINPLWHIRHLGWNPDARYSEHFLQEAKLLHWN  
GRHKPWDFPSVHNDLWESWFVPDPAGIFKLNHHS

**Signal peptide:**

amino acids 1-20

**N-glycosylation site.**

amino acids 234-238

**Tyrosine kinase phosphorylation site.**

amino acids 253-261

**N-myristoylation sites.**

amino acids 63-69, 86-92, 198-204, 218-224, 229-235, 265-271,  
266-272



## **FIGURE 172**

MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLVPQVRTSYNFGRTFLGLDKCNAC  
IGTSICKKFFKEEIRSDNWLASHLGLPPDSLLSYPANYSDDSKIWRPVEIFRLVSKYQNEISD  
RRICASASAPKTCSIERVLRKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCMLR

### **Signal peptide:**

amino acids 1-28

### **N-glycosylation site.**

amino acids 100-103

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 158-161

### **N-myristoylation sites.**

amino acids 56-61, 65-70

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 18-28

### **Prenyl group binding site (CAAX box).**

amino acids 179-182

### **Leucine zipper pattern.**

amino acids 5-26



**FIGURE 173**

GCTGGACTGCTCGCTGGCCGGCAGCGCACCGTTTTGAAGGTCCTAGCCCACCTGGGCTGGCTC  
ACGCGCACGACTAGCCGCTCCCATACAGCACGCCCGGACTCTGTCTGCTCGCTTAAGGCCACTCC  
TATTCTACGGCTGACCCCTGGTGGTCACGTGGATCTGTTCCGCCACGCAAGTCTGGGTCCTTCG  
GCGATTGACCGGGGTCCTTGCTGTTCTGGGAGCCTCTCCTAAGCTGCCTGTTCTGCGCGAGAGTT  
TGGAGGGGCGGGTTTGGGGTCTGGTGTCTGATTGGGGCTCGCACCGCAGCACGCTGGAGTCCCC  
CTTAGGTACCAGTTAGCGTCAGGGGAGCTGGGTCAGGCGGTCTGCCGGGACACCCCGTGTGTGG  
CAGGCGGCGAAGCGCTCTGGAGAATCCCGGACAGCCCTGCTCCCTGCAGCCAGGTGTAGTTTC  
GGGAGCCACTGGGGCCAAAGTGAGAGTCCAGCGGTCTTCCAGCGCTTGGGGCCACGGCGGCGGG  
CCTGGGAGCAGAGGTGGAGCGACCCATTACGCTAAAGATGAAAGGCTGGGGTTGGCTGGCCC  
TGCTTCTGGGGGCCCTGCTGGGAACCGCCTGGGCTCGGAGGAGCCAGGATCTCCACTGTGGAG  
CATGCAGGGCTCTGGTGGATGAACTAGAAATGGGAAATTGCCCAGGTGGACCCCAAGAAGACCA  
TTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAGTGGTGGAGGTGCCTTATG  
CCCGCTCAGAGGCCACCTCACAGAGCTGCTGGAGGAGATATGTGACCGGATGAAGGAGTATG  
GGGAACAGATTGATCCTTCCACCCATCGCAAGAACTACGTACGTGTAGTGGGCCGGAATGGAG  
AATCCAGTGAACCTGGACCTACAAGGCATCCGAATCGACTCAGATATTAGCGGCACCCCTCAAGT  
TTGCGTGTGAGAGCATTGTGGAGGAATACGAGGATGAACTCATTGAATTCTTTTCCCGAGAGG  
CTGACAATGTTAAAGACAACTTTGCAGTAAGCGAACAGATCTTTGTGACCATGCCCTGCACA  
TATCGCATGATGAGCTATGAACCACTGGAGCAGCCCACACTGGCTTGATGGATCACCCCCAGG  
AGGGGAAAATGGTGGCAATGCCTTTTATATATTATGTTTTTACTGAAATTAAGTAAAAATA  
TGAAACCAAAAGT

## **FIGURE 174**

MKGWGWLALLLGALLGTAWARRSQDLHCGACRALVDELEWEIAQVDPKKTIQMGsFRINPDGS  
QSVVEVPYARSEAHLTELEEEICDRMKEYGEQIDPSTHRKNYVRVVGRNGESSELDLQGIRID  
SDISGTLKFACESIVEEYEDELIEFFSREADNVKDKLCSKRTDLCDHALHISHDEL

### **Signal peptide:**

amino acids 1-20

### **N-myristoylation sites.**

amino acids 12-18, 16-22, 29-35

### **Endoplasmic reticulum targeting sequence.**

amino acids 179-184

FIGURE 175

CGCAGCGCGGCAGTCCCTGATGGGCCCCGGCATGGGGTTACCGCTGCTGCCCCCTGCTGTCTGCTCCTGCTGTCGGCGCGTGGCTCAAGCTAGGAAATGGACAGGGCTACTAGCATGGTCCAACTGCAGGGTGGGAGATTCCTGATGGGAACAAATTTCTCCAGACAGCAGAGATGGTGAAGGGCCTGTGCGGGAGGCGACAGTGAAACCCTTTGCCATCGACATATTTCTGTACCAACAAAGATTTTCAGGGATTTTGTACGGGAGAAAAAGTATCGGACAGAAGCTGAGATGTTTGGATGGAGCTTTGTCTTTGAGGACTTTGTCTCTGATGAGCTGAGAAACAAAGCCACCCAGCCAATGAAGTCTGTACTCTGGTGGCTTCCAGTGGAAGAGGCATTTTGGAGGCAGCCTGCAGGTCCTGGCTCTGGCATCCGAGAGAGACTGGAGACCCAGTGTTACACGTGAGCTGGAATGACGCCCCGTGCCTACTGTGCTTGGCGGGGAAAACGACTGCCACGGAGGAAGAGTGGGAGTTTGCCGCCCCGAGGGGGCTTGAAGGGTCAAGTTTACCCATGGGGGAAGTGGTTCCAGCCAAACCGCACCAACCTGTGGCAGGGAAAGTTCCCCAAGGGAGACAAAGCTGAGGATGGCTTCCATGGAGTCTCCCCAGTGAATGCTTTCCCCGCCCAGAACAACTACGGGCTCTATGACCTCCTGGGGAACGTGTGGGAGTGGACAGCATCACCGTACCAGGCTGCTGAGCAGGACATGCGCGTCCTCCGGGGGGGCATCCTGGATCGACACAGCTGATGGCTCTGCCAATCACCGGGCCCCGGGTCACCACCAGGATGGGCAACACTCCAGATTCAGCCTCAGACAACCTCGGTTTCGCTGTGCTGCAGACGCAGGCCGGCCGCCAGGGGAGCTGTAAGCAGCCGGGTGGTGACAAGGAGAAAAGCCTTCTAGGGTCACTGTCATTCCCTGGCCATGTTGCAAACAGCGCAATTCCAAGCTCGAGAGCTTCAGCCTCAGGAAAGAACTTCCCCTTCCCTGTCTCCCATCCCTCTGTGGCAGGCGCTCTCACCAAGGGCAGGAGAGGACTCAGCCTCCTGTGTTTTGGAGAAGGGGCCCAATGTGTGTTGACGATGGCTGGGGGCCAGGTGTTTCTGTTAGAGGCCAAGTATTATTGACACAGGATTGCAAAACACACAAACAGTTGGAACAGAGCACTCTGAAAGGCCATTTTTTTAAGCATTTTAAAATCTATTCCTCTCCCCCTTTCTCCCTGGATGATTCAGGAAGCTGACATTGTTTCCTCAAGGCAGAATTTTCTGGTTCTGTTTTCTCAGCCAGTTGCTGTGGAAGGAGAATGCTTTCTTTGTGGCCTCATCTGTGTTTCGTGTCCCTCTGAAGGAACTAGTTTCCACTGTGTAACAGGCAGACATGTAACATTTTAAGCACAGTTCAGTCCTAAAAGGGTCTGGGAGAACCAGATGATGTACTAGGTGAAGCATTGCAATTGTGGGAATCACAAAGCAAATAGTACTCCAGAAAGACAAATATCAGAAGCTTCCTATTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGGTCTTTCTCTGTTGCCCAGGCTAGAGTGCACTGTGATCACGGCTCACTCTAGCCTTGAATTCCCTGGGCCCAAGCAATTCTCCACCTCAGCCTCTTGAGTAGCTGGGACTACAAGTGTGCACCACCATGCCTGGCTAATTTTTTTGAATTTTTGTAGTGATGGGATCTCGCTCTGTTGCCCAGGGTGGTCTCGAACTCCTGGCCTCAAGCGATCCTCCCACCTCGACCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCTCGCCTGGGCCCCCTTCTCCATATGCCTCCAAAAACATGTCCCTGGAGAGTAGCCTGCTCCACACTGTCACTGGATGTCATGGGGCCAATAAAATCTCCTGCAATTGTGTATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 176**

MARHGLPLLPLLSLLVGAWLKLGNQGQATSMVQLQGGRFLMGNTNSPDSRDGEGPVREATVKPFA  
IDIFPVTNKDFRDFVREKKYRTEAEMFGWSEVFEDEVSEDLRNKATQPMKSVLWWLPVEKAFW  
RQPAGPGSGIRERLEHPVLHVSWNDARAYCAWRGKRLPTEEEWEFAARGGLKGQVYPWGNWFQ  
PNRTNLWQGKFPGDKAEDGFHGVSPVNAFPAQNNYGLYDLLGNVWEWTASPYQAAEQDMRVL  
RGASWIDTADGSANHRARVTTRMGNTPDASDNLGFRCAADAGRPPGEL

**Signal peptide:**

amino acids 1-20

**N-glycosylation site.**

amino acids 191-195

**N-myristoylation sites.**

amino acids 23-29, 25-31, 175-181

**Amidation site.**

amino acids 159-163

FIGURE 177

*(continued)*

## **FIGURE 178**

MPCAQRSWLANLSVVAQLLNFGALCYGRQPQPGPVRFPD RRQEHFIKGLPEYHVVG PVRVDAS  
GHFLSYGLHY PITSSRRKRDLDGSEDWVYYRISHEEKDLFFNLTVNQGF LSNSYIMEKRYGNL  
SHVKMMASSAPLCHLSGTVLQQGTRVGTAALSACHGLTGFFQLPHG DFFIEPVKKHPLVEGGY  
HPHIVYRRQKVPETKEPTCGLKGIVTHMSSWVEESVLFFW

### **Signal peptide:**

amino acids 1-27

### **N-glycosylation sites.**

amino acids 11-15, 105-109, 125-129

### **N-myristoylation site.**

amino acids 149-155

**FIGURE 179**

CAGATTATAAAAGAAACCTTTACTGAATCAGCTGAGTGTTAATAATACGAATTTCTTTTCT  
TGCCAATTCTGATCTGAACAGAAAATCCAAGAACAGGGATATGTGTGGATTACAGTTTTCTCT  
GCCTTGCCTACGACTGTTTCTGGTTGTTACCTGTTATCTTTATTATTACTCCACAAAGAAAT  
ACTTGGAATGTTTCGTCTGTTTGTGACCTCTGCCTGGGAGACAAATTAAGTCCGTAAGTCTAGG  
CCTTTGAGTATTCCTAAGAATTTTCTGAAAGTACAGTTTTTCTGTATCTGACTGGGAATAA  
TATATCTTATATAAATGAAAGTGAATTAACAGGACTTCATTCTCTTGTAGCATTGTATTTGGA  
TAATTCTAACATTCTGTATGTATATCCAAAAGCCTTTGTTCAATTGAGGCATCTATATTTTCT  
ATTTCTAATAATAAATTTTCATCAAACGCTTAGATCCTGGAATATTTAAGGGACTTTTAAATCT  
TCGTAATTTTATATTTACAGTATAATCAGGTATCTTTTGTTCGAGAGGAGTATTTAATGATCT  
AGTTTCAGTTCAGTACTTAAATCTACAAAGGAATCGCCTCACTGTCCTTGGGAGTGGTACCTT  
TGTTGGTATGGTTGCTCTTCGGATACTTGATTTATCAAACAATAACATTTTGAGGATATCAGA  
ATCAGGCTTTCAACATCTTGAAAACCTTGCTTGTTGTATTTAGGAAGTAATAATTTAACAAA  
AGTACCATCAAATGCCTTTGAAGTACTTAAAGTCTTAGAAGACTTTCTTTGTCTCATAATCC  
TATTGAAGCAATACAGCCCTTTGCATTTAAAGGACTTGCCAATCTGGAATACCTCCTCCTGAA  
AAATTCAGAATTAGGAATGTTACTAGGGATGGGTTTAGTGGAATTAATAATCTTAAACATTT  
GATCTTAAGTCATAATGATTTAGAGAATTTAAATTTCTGACACATTCAGTTTGTTAAAGAATTT  
AATTTACCTTAAGTTAGATAGAAACAGAATAATTAGCATTGATAATGATACATTTGAAAATAT  
GGGAGCATCTTTGAAGATCCTTAATCTGTCATTTAATAATCTTACAGCCTTGCATCCAAGGGT  
CCTTAAGCCGTTGTCTTCATTGATTTCATCTTCAGGCAAATTTCTAATCCTTGGGAATGTAAGT  
CAAACTTTTGGGCCTTCGAGACTGGCTAGCATCTTCAGCCATTACTCTAAACATCTATTGTCA  
GAATCCCCCATCCATGCGTGGCAGAGCATTACGTTATATTAACATTACAAATTGTGTTACATC  
TTCAATAAATGTATCCAGAGCTTGGGCTGTTGTAAATCTCCTCATATTCATCACAAGACTAC  
TGCGCTAATGATGGCCTGGCATAAAGTAACCACAAATGGCAGTCCTCTGGAAAATACTGAGAC  
TGAGAACATTACTTTCTGGGAACGAATTCCTACTTCACCTGCTGGTAGATTTTTTTCAAGAGAA  
TGCCTTTGGTAATCCATTAGAGACTACAGCAGTGTTACCTGTGCAAATACAACCTTACTACTTC  
TGTTACCTTGAACCTTGGAACAAAACAGTGCTCTACCGAATGATGCTGCTTCAATGTCAGGGAA  
AACATCTCTAATTTGTACACAAGAAGTTGAGAAGTTGAATGAGGCTTTTGACATTTTGCTAGC  
TTTTTTCATCTTAGCTTGTGTTTTAATCATTTTTTTGATCTACAAAGTTGTTCAAGTTTAAACA  
AAAACCTAAAGGCATCAGAAAACCTCAAGGGAAAATAGACTTGAATACTACAGCTTTTATCAGTC  
AGCAAGGTATAATGTAAGTGCCTCAATTTGTAACACTTCCCCAAATTTCTCTAGAAAGTCTGG  
CTTGGAGCAGATTCGACTTCATAAACAAATTGTTTCTGAAAATGAGGCACAGGTCATTCTTTT  
TGAACATTCTGCTTTATAACTCAACTAAATATTGTCTATAAGAACTTCAGTGCCATGGACAT  
GATTTAAACTGAAACCTCCTTATATAATTATATACTTTAGTTGGAAATATAATGAATTATATG  
AGGTTAGCATTATTAAAAATATGTTTTTTNTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 180**

MCGLQFSLPCLRLFLVVTCYLLLLLHKEILGCSSVCQLCTGRQINCRNLGLSSI PKNFPESTV  
 FLYLTGNNISYINESELTGLHSLVALYLDNSNILYVYPKAFVQLRHLYFLFLNNNFIKRLDPG  
 IFKGLLNLRNLYLQYNQVSFVPRGVFNDLVSQYLNQLQRNRLTVLGSGTFVGMVALRILDLSN  
 NNILRISESGFQHLENLACLYLGSNNLT KVPSNAFEVLKSLRRLSLSHNPIEAIQPF AFKGLA  
 NLEYLLLKNSRIRNVTRDGFSGINNLKHLILSHNDLENLNSDTFSLLKNLIY LKLDNRNIISI  
 DNDTFENMGASLKILNLSFNNTALHPRVLKPLSSLIHLQANSNPWECNCKLLGLRDWLASSA  
 ITLNIYCQNPPSMRGRALRYINITNCVTSSINVSRAWAVVKSPHIHHKTTALMMAWHKVT TNG  
 SPLENTETENITFWERIPTSPAGRFFQENAFGNPLETTAVLPVQIQLTTSVTLNLEKNSALPN  
 DAASMSGKTSLIC TQEVEKLNEAFDILLAFFILACVLIIFLIYKVVQFKQKLKASENSRENRL  
 EYYSFYQSARYNVTASICNTSPNSLES PGLEQIRLHKQIVPENE AQVILFEHSAL

**Signal peptide:**

amino acids 1-41

**Transmembrane domain:**

amino acids 530-547

**N-glycosylation sites.**

amino acids 71-75, 76-80, 215-219, 266-270, 317-321, 331-335,  
 336-340, 400-404, 410-414, 451-455, 579-583

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 231-235

**N-myristoylation sites.**

amino acids 3-9, 69-75, 126-132, 174-180

**ATP/GTP-binding site motif A (P-loop).**

amino acids 506-514



**FIGURE 181**

GGCCTGGCGCGGCGCTCCGGTAAGGCGTGTGTGCGGCAGGGCGGGGACAGAACCGTCCTCTCG  
GGCTCTGGGCGTGTCCGAGACCGCGCTCCCCGCCGAAATCAAGCTCCGAGTCATCCGTGTGGG  
GCATTCGTCCCCCCTGGCACAGTTGGCCTCTTTCCAGAAGCCCGTTTTGTTTGTTTTACGTCT  
AAATTCGCGTCGGTCTTATTTCTCTCCCTGGCAAGGTCTGAAGACGGGTAGGAGAATAACCT  
GTGTCAGCGTGTT**ATG**ATGCCGTCCCGTACCAACCTGGCTACTGGAATCCCCAGTAGTAAAGT  
GAAATATTCAAGGCTCTCCAGCACAGACGATGGCTACATTGACCTTCAGTTTAAGAAAACCCC  
TCCTAAGATCCCTTATAAGGCCATCGCACTTGCCACTGTGCTGTTTTTGATTGGCGCCTTTCT  
CATTATTATAGGCTCCCTCCTGCTGTCAGGCTACATCAGCAAAGGGGGGGCAGACCGGGCCGT  
TCCAGTGCTGATCATTGGCATTCTGGTGTTCCTACCCGGATTTTACCACCTGCGCATCGCTTA  
CTATGCATCCAAAGGCTACCGTGGTTACTCCTATGATGACATTCCAGACTTTGATGAC**TAG**CA  
CCCACCCCATAGCTGAGGAGGAGTCACAGTGGAAGTGTCCAGCTTTAAGATATCTAGCAGAA  
ACTATAGCTGAGGACTAAGGAATTCTGCAGCTTGCAGATGTTTAAGAAAATAATGGCCAGATT  
TTTTGGGTCCCTCCCAAAGATGTTAAGTGAACCTACAGTTAGCTAATTAGGACAAGCTCTATT  
TTTCATCCCTGGGCCCTGACAAGTTTTTCCACAGGAATATGTATCATGGAAGAATAGAGGTTA  
TTCTGTAATGGAAAAGTGTTCCTGCCACCACCCTCTGTAGAGCTGAGCATTTCTTTTAAATA  
GTCTTCATTGCCAATTTGTTCTTGTAGCAAATGGAACAATGTGGTATGGCTAATTTCTTATTA  
TTAAGTAGTTTATTTTAAAAATATCTGAGTATATTATCCTGTACACTTATCCCTACCTTCATG  
TTCCAGTGGAAGACCTTAGTAAAATCAAAGATCAGTGAGTTCATCTGTAATATTTTTTTTACT  
TGCTTTCTTACTGACAGCAACCAGGAATTTTTTTTATCCTGCAGAGCAAGTTTTCAAATGTAA  
ATACTTCCTCTGTTTAAACAGTCCTTGGACCATTCTGATCCAGTTCACCAGTAGGTTGGACAGC  
ATATAATTTGCATCATTTTGTCCCTTGTAATCAAGATGTTCTGCAGATTATTCCTTTAACGG  
CCGGACTTTTGGCTGTTTCCTAATGAAACATGTAGTGGTTATTATTTAGAGTTTATAGCCGTA  
TTGCTAGCACCTTGTAGTATGTCATCATTCTGCTCATGATTCCAAGGATCAGCCTGGATGCCT  
AGAGGACTAGATCACCTTAGTTTGATTCTATTTTTTTAGCTTGCAAAAAGTGACTTATATTCCA  
AAGAAATTAAATGTTGAAATCCAAATCCTAGAAATAAAATGAGTTTNNTTCCAAAAAAAAAA  
AA

## **FIGURE 182**

MMPSRTNLATGIPSSKVKYSRLSSTDDGYIDLQFKKTPPKIPYKAIALATVLFLIGAFLIIIG  
SLLLSGYISKGGADRAVPVLIIGILVFLPGFYHLRIAYYASKGYRGYSYDDIPDFDD

**Transmembrane domains:**

amino acids 45-66, 79-95

**N-myristoylation sites.**

amino acids 11-17, 75-81

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 183**

CTAAAAAATACAAAAATTAGCTGGGCGTGGTGTCTATGTACCTGTAATCCCAGCTACTCAAGAGGCTGAGGCAGGA  
 GAATCGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCAAGATTAAGTCACTGCCTCCAGCCTGGGTGACAGA  
 GCAAGACTCTGTATCAAAATAAAATAAATAAAGTACAACCTCTGGATGGGCATGGTGGCTTATGTCTGTAATCCCAG  
 CACTTTGGGAACCTTGAGGCGGGTAGATTGCTTGAGTCCGGGAGTTTGAGACCAGTCTGGGTAATATGGTAACCCCT  
 GTCTACCAAAAAATACAGGTATTAGCCAGTCTCATAACTCGGTCTCAAAATAAATAAATACATACATACATAGATG  
 AAAATTTAAAAAATAAAGTCCAACCTCAGCGGTTTTTCAGCATATTTACAGAGTTGTACAATCTTCACCACTATCTA  
 ATTTCAGAACATTTTCATCACCCCCAAAAGAAACCTAACCCATTGACTATCTCTCCATTTCCCTCCCTCTCCCTAG  
 CCTCTGGCAACCCTAATCTCTTTTTTGTCTCTATAGATTTGCCTATTTTGGACAGTTCATATACAAGGAATCAT  
 ACCACATGTAGCCTTTTGTGTCCGGCTTCTTTGATTAATAGAATGTTTTCAAGGCTCATCTATGCTGTAGCCTGT  
 ATCAGCACTTCATTCTCTTCTATGGCTGAATAATAGTCCACTGTAGGGATGTGCCATGTTTTTCCACTAGCTGAT  
 GGACATTTGGGTTGTTTCCACCTTCTGGCTATTATAAATATTGCTGCTATAAATATTCACTTACAAGTTTTTGTG  
 TGGACATATGTTTTTATTTCTTCTGGTATATCCTTCGGAGTGGAACCTGCTGGATCAGGTGGTAACCTTAGGTCTA  
 ACCTGGCAGTTAAACAGAATCCTATGCATGCTGTAGTCCATGAGTTGAAATAAACACTTGACCCATAGTAAGTGC  
 CAGATCATCTTCATTTACAGCAACCAGTAATTTACAGATGAGGAAATGAAGGCTCCAGAGGTGAACTGGCTT  
 TTCCCATTTGAGCAGTTCCAAGTCAGACAGTTAAAAAGTGGCAGGACCTGGAAGAGAAGCTAGTTCTTTACCCCT  
 GGCATTCAGGGCTGCCTCCTGGGCTACGGGGCTGGCATTTAGAATAGAGCTAAGGTCTGCTGCCAAGGCAGGTGC  
 CCCAGTCTGCCTCCTCTGTGTCTTATTCCACTTTCTCTGCAGCCCTCCAGGGGACCCCTCTCTCAGCCACCCCTC  
 TCTCTGGT**GATGT**CACAGTGCTGCCGGAAGATCAAAGATACGGTGCAGAACTGGCTTCGGACCATAGGACATT  
 CACAGCAGTGATCCCCGAGTGGGCAAAGCCATTGACAGGAACTTCGACTCTGAGATCTGTGGTGTGTGTGCAGAT  
 GCGGTGTGGGACGCGCGGGAACAGCAGCAGCAGATCCTGCAGATGGCCATCGTGGAACACCTGTATCAGCAGGGC  
 ATGCTCAGCGTGGCCGAGGAGCTGTGCCAGGAATCAACGCTGAATGTGGACTTGGATTTCAAGCAGCCTTTCTTA  
 GAGTTGAATCGAATCCTGGAAGCCCTGCACGAACAAGACCTGGGTCTGCGTTGGAATGGGCCGTCTCCACAGG  
 CAGCGCCTGCTGGAACCTCAACAGCTCCCTGGAGTTCAAGCTGCACCGACTGCACCTCATCCGCCCTCTTGGCAGGA  
 GGCCCCGCGAAGCAGCTGGAGGCCCTCAGCTATGCTCGGCACTTCCAGCCCTTTGCTCGGCTGCACCAGCGGGAG  
 ATCCAGGTGATGATGGGCAGCCTGGTGTACCTGCGGCTGGGCTTGGAGAACTCACCCCTACTGCCACCTGCTGGAC  
 AGCAGCCACTGGGCAGAGATCTGTGAGACCTTTACCCGGGACGCTGTTCCTGCTGGGGCTTTCTGTGGAGTCC  
 CCCCTTAGCGTCAGCTTTGCCTCTGGCTGTGTGGCGCTGCCTGTGTTGATGAACATCAAGGCTGTGATTGAGCAG  
 CGGCAGTGCCTGGGGTCTGGAATCACAAGGACGAGTTACCGATTGAGATTGAACTAGGCATGAAGTGCTGGTAC  
 GCTCATCTGTGGCCATGTTATCTCCCAGATGCACCTCAATAAGCTCATTAAATGGAGGAAACACTCCGTGTTGCT  
 TGCCCCATCCTCCGCCAGCAGACGTCAGATTCCAACCTCCCATCAAGCTGAAGTGTCCTACTGTCCCATGGAG  
 CAGAACCCGGCAGATGGGAAACGCATCATATCT**TGA**TTCTACCTGGAAGGAATTTTGTGAAAGGGGTTTTTCAC  
 CTGTGAGCCTTGGTCTGTCTCGGTAGGGTGGTCAACTTCAGTGGACTGTGGTTGGTTTCAGAGCGCCTGGCTGAG  
 GAGTTCCACTGAGGGGAGCACTGGAGCAGCCCTTTGGCAGAGGCTGAGGAGGGAGATGGACCAGCCCACGCCTGG  
 CACCTGGCTCCATGGCATAAGGAAAGGGAGATGCTGGCCTCTGTGCTCCTGCTGTCTTTTCTGTTTCTGTTTGC  
 GTTTGACTTAGTAGCAACCGACAGAGTGGCAAGGGATTGGTCTTCAGCAGTAGACATCCTTCCACCCCTGCCCT  
 CAGCCAAGTCTCTTGCTGCCATGCCAATGCTATGTCCACCCCTGCCCTCGGCCCCAAGAGTGTCCAGCGGTGGCC  
 CACCTCTTCTCTCCACTACAGCCTCAACAGTATGTACCATCTCCACTGTAAATAGTCCCAGTTAGAACGGAATG  
 CCGTTGTTTTATAACTTTGAACAAATGTATTTACTGCCCTTCTCAAAA

10028072 = 131901

## **FIGURE 184**

QCCRKIKDTVQKLASDHKDIHSSVSRVGKAIDRNFSEICGVVSDAVWDAREQQQQILQMAIV  
EHLYQQGMLSVAEELCQESTLNVDLDFKQPFLELNRIEALHEQDLGPALEWAVSHRQRLLLEL  
NSSLEFKLHRLHFIRLLAGGPAKQLEALSYARHFQPFARLHQREIQVMMGSLVYLRLGLEKSP  
YCHLLDSSHWAEICETFTRDACSLGLSVESPLSVSFASGCVALPVLMMNIKAVIEQRQCTGVW  
NHKDELPIEIELGMKCWYHSVFACPILRQQTSDSNPPIKLCGHVISRDALNKLINGGKLKCP  
YCPMEQNPADGKRIIF

### **Transmembrane domain:**

amino acids 222-241

### **N-glycosylation site.**

amino acids 129-133

### **Tyrosine kinase phosphorylation site.**

amino acids 151-159, 184-193

### **Amidation site.**

amino acids 327-331

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 222-233

**FIGURE 185**

GAGCGACGCTGTCTCTAGTCGCTGATCCCAA**ATG**CACCGGCTCATCTTTGTCTACACTCTAAT  
CTGCGCAAACCTTTTGCAGCTGTCGGGACACTTCTGCAACCCCGCAGAGCGCATCCATCAAAGC  
TTTGCGCAACGCCAACCTCAGGCGAGATGACTTGTACCGAAGAGATGAGACCATCCAGGTGAA  
AGGAAACGGCTACGTGCAGAGTCCTAGATTCCCGAACAGCTACCCAGGAACCTGCTCCTGAC  
ATGGCGGCTTCACTCTCAGGAGAATACACGGATACAGCTAGTGTTTGACAATCAGTTTGGATT  
AGAGGAAGCAGAAAATGATATCTGTAGGTATGATTTTGTGGAAGTTGAAGATATATCCGAAAC  
CAGTACCATTATTAGAGGACGATGGTGTGGACACAAGGAAGTTCCTCCAAGGATAAAATCAAG  
AACGAACCAAATTAAAATCACATTCAAGTCCGATGACTACTTTGTGGCTAAACCTGGATTCAA  
GATTTATTATTCTTTGCTGGAAGATTTCCAACCCGCAGCAGCTTCAGAGACCAACTGGGAATC  
TGTCACAAGCTCTATTTCAGGGGTATCCTATAACTCTCCATCAGTAACGGATCCCCTCTGAT  
TGCGGATGCTCTGGACAAAAAAATTGCAGAATTTGATACAGTGGAAGATCTGCTCAAGTACTT  
CAATCCAGAGTCATGGCAAGAAGATCTTGAGAATATGTATCTGGACACCCCTCGGTATCGAGG  
CAGGTCATACCATGACCGGAAGTCAAAGTTGACCTGGATAGGCTCAATGATGATGCCAAGCG  
TTACAGTTGCACTCCCAGGAATTACTCGGTCAATATAAGAGAAGAGCTGAAGTTGGCCAATGT  
GGTCTTCTTTCCACGTTGCCTCCTCGTGCAGCGCTGTGGAGGAAATTGTGGCTGTGGAAGTGT  
CAACTGGAGGTCCTGCACATGCAATTGAGGGAAAACCGTGAAAAAGTATCATGAGGTATTACA  
GTTTGAGCCTGGCCACATCAAGAGGAGGGGTAGAGCTAAGACCATGGCTCTAGTTGACATCCA  
GTTGGATCACCATGAACGATGCGATTGTATCTGCAGCTCAAGACCACCTCGAT**TAA**GAGAATGT  
GCACATCCTTACATTAAGCCTGAGAGAA

**FIGURE 186**

MHRLIFVYTLICANFCSCRDTSATPQSASIKALRNANLRRDDLYRRDETIQVKNGYVQSPRF  
PNSYPRNLLLTWRLHSQENTRIQLVFDNQFGLLEEAENDICRYDFVEVEDISETSTIIRGRWCG  
HKEVPPRIKSRTNQIKITFKSDDYFVAKPGFKIYYSLLEDFQPAAASETNWESVTSSISGVSY  
NSPSVTDPTLIADALDKKIAEFDTVEDLLKYFNPESWQEDLENMYLDTPRYRGRSYHDRKSKV  
DLDRLNDDAKRYSCTPRNYSVNI REELKLANVVFPRCLLVQRCGGNCGCGTVNWRSCTCNSG  
KTVKKYHEVLQFEPGHIKRRGRAKTMALVDIQLDHHERCDCICSSRPPR

Signal peptide:

amino acids 1-18

N-glycosylation site.

amino acids 270-274

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 262-266

Tyrosine kinase phosphorylation site.

amino acids 256-265

N-myristoylation sites.

amino acids 94-100, 186-192, 297-303, 298-304

TonB-dependent receptor proteins signature 1.

amino acids 1-56

**FIGURE 187**

**CATG**CCGCTGCCGCCGCTGCTGCTGTTGCTCTCCTGGCGGCGCCTTGGGGACGGCAGTTCCCTG  
TGTCTCTGGTGGTTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGT  
CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTT  
CATATATGGGCAAAGAAATGGCTGAATAAATCAGAATGCAGAAATATCAATAGAACCTACTG  
TGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTG  
GGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCTATCCTTTTTTTAGAAACACAAAT  
TGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCTTGACAGCTCC  
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCT  
GAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAA  
CCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTT  
CGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGA  
TCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGT  
GTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACA  
CCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGA  
AAAAATCGTGATTAACTTTATCACCCCTCAATATCTCGGATGATTCTAAATTTCTCATCAGGA  
TATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAA  
CCTGAGGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGA  
AATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCAGGAGCAAGAGTCCCTCAG  
CAGAACAATACCCCCGGATAAAAACAGTCAATTGAATATGAATATGATGTCAGAACCACTGACAT  
TTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATT  
ATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCCGCAAACGTTACAGTACTCATAACCCCC  
TCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCGGAGGA  
AGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAACTGGCAGGCTGTGTATTCTTTCGCT  
GTCCAGCTTCGACCAGGATTGAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGA  
GGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAAC  
CTATCTCATGCAATTCATGGAGGAATGGGGGTTATATGTGCAGATGGAAAAC**TGA**TGCCAACA  
CTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAA  
GTACCTGGGATGAAAGAAGTTTTTTTCCAGTTTGTGAGTGTCTGTGAGAA

**FIGURE 188**

MPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGVKVTYTVQYF  
IYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPFLETQI  
GPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTWSQCVTN  
HTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLKDQSSEFKAKIIFWYVLPISITV  
FLFSVMGYSIYRYIHVGKEKHPANLILYGNEDKRFVPAEKIVINFITLNISSDDSKISHQD  
MSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENTEGTSLTQQESLS  
RTIPDPKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLSQALAVLGPQTLQYSYTP  
QLQDLPLAQEHTDSEEGPEEEPPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEE  
GLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

**Signal sequence:**

amino acids 1-18

**Transmembrane domain:**

amino acids 240-260

**N-glycosylation sites.**

amino acids 31-34, 72-75, 80-83, 171-174, 180-183, 189-192,  
304-307, 523-526

**Tyrosine kinase phosphorylation site.**

amino acids 385-392, 518-526

**N-myristoylation sites.**

amino acids 53-58, 106-111, 368-373, 492-497

**Tissue factor**

amino acids 1-278



**FIGURE 189**

**ATG**TGCTGCTGGCCGCTGCTCCTGCTGTGGGGGCTGCTCCCCGGGACGGCGGGCGGGGGGCTCG  
GGCCGAACCTATCCGCACCGGACCTCCTGGACTCGGAGGGCAAGTACTGGCTGGGCTGGAGC  
CAGCGGGGCAGCCAGATCGCCTTCCGCCTCCAGGTGCGCACTGCAGGCTACGTGGGCTTCGGC  
TTCTCGCCACCGGGGCCATGGCGTCCGCCGACATCGTCGTGGGCGGGGTGGCCCACGGGCGG  
CCCTACCTCCAGGATTATTTTACAAATGCAAATAGAGAGTTGAAAAAGATGCTCAGCAAGAT  
TACCATCTAGAATATGCCATGGAAAATAGCACACACACAATAATTGAATTTACCAGAGAGCTG  
CATACATGTGACATAAATGACAAGAGTATAACGGATAGCACTGTGAGAGTGATCTGGGCCTAC  
CACCATGAAGATGCAGGAGAAGCTGGTCCCAAGTACCATGACTCCAATAGGGGCACCAAGAGT  
TTGCGGTTATTGAATCCTGAGAAAAGTAGTGTGCTATCTACAGCCTTACCATACTTTGATCTG  
GTAAATCAGGACGTCCCCATCCCAAACAAAGATACAACATATTGGTGCCAAATGTTTAAGATT  
CCTGTGTTCCAAGAAAAGCATCATGTAATAAAGGTTGAGCCAGTGATACAGAGAGGCCATGAG  
AGTCTGGTGCACCACATCCTGCTCTATCAGTGCAGCAACAACTTTAACGACAGCGTTCTGGAG  
TCCGGCCACGAGTGCTATCACCCCAACATGCCCGATGCATTCCCTCACCTGTGAAACTGTGATT  
TTTGCCTGGGCTATTGGTGGAGAGGGCTTTTCTTATCCACCTCATGTTGGATTATCCCTTGGC  
ACTCCATTAGATCCGCATTATGTGCTCCTAGAAGTCCATTATGATAATCCCACTTATGAGGAA  
GGCTTAATAGATAATTCTGGACTGAGGTTATTTTACACAATGGATATAAGGAAATATGATGCT  
GGGGTGATTGAGGCTGGCCTCTGGGTGAGCCTCTTCCATACCATCCCTCCAGGGATGCCTGAG  
TTCCAGTCTGAGGGTCACTGCACTTTGGAGTGCCTGGAAGAGGCTCTGGAAGCCGAAAAGCCA  
AGTGGAATTCATGTGTTTGCTGTTCTTCTCCATGCTCACCTGGCTGGCAGAGGCATCAGGCTG  
CGTCATTTTCGAAAAGGGAAGGAAATGAAATTACTTGCCTATGATGATGATTTTGACTTCAAT  
TTCCAGGAGTTTCAGTATCTAAAGGAAGAACAACAATCTTACCAGGAGATAACCTAATTACT  
GAGTGTGCTACAACACGAAAGATAGAGCTGAGATGACTTGGGGAGGACTAAGCACCAGGAGT  
GAAATGTGTCTCTCATACCTTCTTTATTACCCAAGAATTAATCTTACTCGATGTGCAAGTATT  
CCAGACATTATGGAACAACCTTCAGTTCATTGGGGTTAAGGAGATCTACAGACCAGTCACGACC  
TGGCCTTTTATTATCAAAAAGTCCCAAGCAATATAAAAACCTTTTCTTTCATGGATGCTATGAAT  
AAGTTTAAATGGACTAAAAAGGAAGGTCTCTCCTTCAACAAGCTGGTCCTCAGCCTGCCAGTG  
AATGTGAGATGTTCCAAGACAGACAATGCTGAGTGGTTCGATTCAAGGAATGACAGCATTACCT  
CCAGATATAGAAAGACCCTATAAAGCAGAACCTTTGGTGTGTGGCACGTCTTCTTCTCTTCC  
CTGCACAGAGATTTCTCCATCAACTTGCTTGTTTGCCTTCTGCTACTCAGCTGCACGCTGAGC  
ACCAAGAGCTTG**TGA**TCAAAAATTCTGTTGGACTTGACAATGTTTTCTATGATCTGAACCTGTC  
ATTTGAAGTACAGGTTAAAGACTGTGTCCACTTTGGGCATGAAGAGTGTGGAGACTTTTCTTC  
CCCATTTTCCCTCCCTCCTTTTTCTTTCCATGTTACATGAGAGACATCAATCAGGTTCTCTT  
CTCTTTCTTAGAAATACCTGATGTTATATATACATGGTCAATAAAAATAAACTGGCCTGACTT  
AAGATAACCATTTTAAAAAATTGGGCTGTCATGTGGGAATAAAAGAATTCTTTCTTTCCTAAA  
AAAAAAA

**FIGURE 190**

MCCWPLLLLWGLLPGTAAGGSGRTYPHRTLDDSEGKYWLGWSQRGSQIAFRLQVVRTAGYVGFG  
 FSPTGAMASADIVVGGVAHGRPYLQDYFTNANRELKKDAQQDYHLEYAMENSTHTTIEFTREL  
 HTCDINDKSITDSTVRVIWAYHHEDAGEAGPKYHDSNRGTKSLRLLNPEKTSVLSTALPYFDL  
 VNQDVPIPNKDTTYWCQMFKIPVFQEKHHVIKVEPVIIQRGHESLVHHILLYQCSNNFNDSVLE  
 SGHECYHPNMPPDAFLTCTETVIFAWAIGGEGFSYPPHVGLSLGTPLDPHYVLLEVHYDNPTYEE  
 GLIDNSGLRRLFYTMDIRKYDAGVIEAGLWVSLFHTIPPGMPEFQSEGHCTLECLEEAEAEKPE  
 SGIHVFAVLLHAHLARGIRLRHFRKKGEMKLLAYDDDFDNFQEFQYLKEEQTILPGDNLIT  
 ECRYNTKDRAEMTWGGLSTRSEMCLSYLLYYPRINLTRCASIPDIMEQLQFIGVKEIYRPVTT  
 WPFIIKSPKQYKNLSFMDAMNKFVKWKKEGLSFNKLVLSPVNVRCSTKTDNAEWSIQGMTALP  
 PDIERPYKAEPLVCGTSSSSSLHRDFSINLLVCLLLLSCITLSTKSL

**Signal peptide:**

amino acids 1-18

**Transmembrane domains:**

amino acids 56-73, 378-393, 583-602

**N-glycosylation sites.**

amino acids 114-118, 247-251, 476-480, 517-521

**N-myristoylation sites.**

amino acids 11-17, 15-21, 20-26, 45-51, 68-74, 79-85, 290-296,  
 316-322, 337-343, 342-348, 456-462, 534-540, 582-588

**Copper type II, ascorbate-dependent monooxygenases proteins.**

amino acids 271-321, 422-474



**FIGURE 192**

MKRLLLFLFFITFSSAFPLVRMTENEENMQLAQAYLNQFYSLEIEGNHLVQSKNRSLIDDKI  
 REMQAFFGLTVTGKLDSENTLEIMKTPRCGVPDVGQYGYTLPGWRKYNLTYRIINYTPDMARAA  
 VDEAIQEGLEVWSKVTPKFTKISKGIADIMIAFRTRVHGRCPRYFDGPLGVLGHAFPPGPGL  
 GGDTHFDEEDENWTKDGAGFNFLVAAHEFGHALGLSHSNDQTALMFPNYVSLDPRKYPLSQDD  
 INGIQSIYGGGLPKVPAKPKEPTIPHACDPDLTFDAITTFRREVMFFKGRHLWRIYYDITDVEF  
 ELIASFWPSLPADLQAAYENPRDKILVFKDENFWMIRGYAVLPDYPKSIHTLGFPGRVKKIDA  
 AVCDKTTRKTYFFVGIWCWRFDEMTQTMKGFPQRVVKHFPGISIRVDAAFQYKGFFFSRGS  
 KQFEYNIKTKNITRIMRTNTWFQCKEPKNSSFGFDINKEKAHSGGIKILYHKSLSLFIFGIVH  
 LLKNTSIYQ

**Signal peptide:**

amino acids 1-17

**N-glycosylation sites.**

amino acids 55-59, 110-114, 200-204, 452-456, 470-474, 508-512

**N-myristoylation site.**

amino acids 71-77, 205-211, 223-229

**Hemopexin domain signature.**

amino acids 171-202, 207-238, 318-334

**Neutral zinc metalloproteinases, zinc-binding region signature.**

amino acids 213-223

**Matrixins cysteine switch.**

amino acids 89-97, 207-238

**FIGURE 193**

CACAATCAGGTCCCATTCTATAGATGGGGAAACTGAGGCTTGAGGTCACATAGGCGTCGTTCA  
AGGCTGGTATACCTGCACCCTCTCCCATGTGAACAAC**ATG**GTTCTGGGTAATGGGGGCTGTCA  
TCCAGTCTCCTCCCTGCCCCCTGCTGGTGCACCTTCCTGCCTCTGCTGGTGCACCTTCTGCCCCCT  
ACTGGTATATTTGCTGCCTCTGCTGGGGCGCTTCCTGCCTCGGCTGGTGTATCTCCTGCCCCCT  
GCTGGTGCACCTTCTGCCCCCGCTGATGCACCTTCCTGCCTCTGCTGGTGCACCTTCCTGGCTCT  
GCTGGCACACTTCCTGCCTCTGCTGGTGCACCTTCCTGGCTCTGCTGGCGCACTTTCCTGCCCC  
TGCTGGTGTATTTCTGCCCCCTGCTGGTGTACTTCCTTCCCCTGCTGGTGCACCTTCCTGCCTC  
TGCTGGCGCACTTCTTGCTCTCCAGGCCCTACCT**TAG**CCTCTCCCTCTTATATATGGAAGTCT  
TCCCAGTTCACTGACACTGGTAACAGGGACTCTGCTCTTGGTGTTGCTGTCTGCCCTGGGGAT  
GGGCATCTGTGTCTTCCTTTACTACTGCTGGCTCAGGACCCAGAGCTTTGAAGCATGTCCAGA  
TGCAGGTCCGGGCACCAGAGTCTAAGGAGCCCCTACACCCACCAGGATTTTCCAATAAAGAGA  
TGTTACCA

## **FIGURE 194**

MVLGNGGCHPVSSLPLLVHFLPLLVHFLPLLVYLLPLLGRFLPRLVYLLPLLVHFLPPLMHFL  
PLLVHFLALLAHFLPLLVHFLALLAHFPAPAGVFPAPAGVLPSPAGALPASAGALLASPGPT

### **Signal peptide:**

amino acids 1-39

### **N-myristoylation sites.**

amino acids 4-10, 109-115, 116-122

### **Leucine zipper pattern.**

amino acids 14-36, 16-38, 17-39, 21-43, 24-46, 28-50, 31-53,  
35-57, 38-60, 42-64, 45-67, 49-71, 52-74, 56-78, 59-81, 63-85,  
65-87, 66-88



**FIGURE 196**

MRRLTRRLVLPVFGVLWITVLLFFWVTKRKLEVPTGPEVQTPKPSDADWDDLWDQFDERRYLN  
AKKWRVGDDPYKLYAFNQRESERISSNRAIPDTRHLRCTLVYCTDLPPTSIIITFHNEARST  
LLRTIRSVLNRTPTHIREIILVDDFSNDPDDCKQLIKLPKVKCLRNNERQGLVRSRIRGADI  
AQGTTLTFLDSHCEVNRDWLQPLLHRVKEDYTRVVCVIDIINLDTFTYIESASELRGGFDWS  
LHFQWEQLSPEQKARRLDPTPIRTPIIAGGLFVIDKAWFDYLGKYDMDMDIWGGENFEISFR  
VWMCSSLEIVPCSRVGHVFRKKHPYVFPDGNANTYIKNTKRTAEVWMDEYKQYYYAARPFAL  
ERPFGNVESRLDLRKNLRCQSFKWYLENIYPELSIPKESSIQKGNIRQRQKCLESQRQNNQET  
PNLKLSPCAKVKGEDAKSQVWAFITYTQQILQEELCLSVITLFPGAPVVLVLCKNGDDRQQWTK  
TGSHEHIAHSLCLDMDMFGDGTENGKEIVVNPCESSLMSQHWDMVSS

**Transmembrane domain:**

amino acids 475-493

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 2-6

**Tyrosine kinase phosphorylation sites.**

amino acids 68-75, 401-409

**N-myristoylation sites.**

amino acids 178-184, 186-192, 192-198, 346-352, 383-389, 526-532





**FIGURE 198**

MGEDDAALRAGSRGLSDPWADSVGVRPRTERHIAVHKRLVLFAVSLVALLAVTMLAVLLSL  
 RFDECGASATPGADGGPSGFFPERGGNGSLPGSARRNHHAGGDSWQPEAGGVASPGTTSAPPS  
 EEEREPWEPWTQLRLSGHLKPLHYNLMLTAFMENFTFSGEVNVEIACRNATRYVVLHASRVAV  
 EKVQLAEDRAFGAVPVAGFFLYPQTQVLVVVLNRTLDAQRNYNLKIIYNALIENELLGFFRSS  
 YVLHGERRFLGVTQFSPTHARKAFPCFDEPIYKATFKISIKHQATYLSLSNMPVETSVFEEDG  
 WVTDFHSQTPLMSTYYLAWAICNFTYRETTTKSGVVVRLYARPDAIRRGSGDYALHITKRLIE  
 FYEDYFKVPYSLPKDLLAVPKHPYAAMENWGLSIFVEQRILLDPSVSSISYLLDVTMIVHE  
 ICHQWFGDLVTPVWVEDVWLKEGFAHYFEFVGTDYLYPGWNMEKQRFLTDVLHEVMLLDGLAS  
 SHPVSQEVQLQATDIDRVFDWIAYKKGAALIRMLANFMGHSVFQRLQDYLTIHKYGNAARNDL  
 WNTLSEALKRNGKYVNIQEVMDQWTLQMGYPVITILGNTTAENRIIITQQHFIYDISAKTKAL  
 KLQNNSYLWQIPLTIVVGNRSHVSSEAIIWVSNKSEHHRITYLDKGSWLLGNINQTYFRVNY  
 DLNRWRLIDQLIRNHEVLSVSNRAGLIDDAFSLARAGYLPQNIPIEIIRYLSEEKDFLPWHA  
 ASRALYPLDKLLDRMENYNIFNEYILKQVATTYIKLGWPKNNFNGSLVQASYQHEELRREVIM  
 LACSGNKHCHQQASTLISDWISSNRNRIPLNVRDIVYCTGVSLLEDVWEFIWMKFHSTTAV  
 SEKKILLEALTCSDDRNLLNRLNLSLNSEVVLDQDAIDVIIHVARNPHGRDLAWKFFRDKWK  
 ILNTRYGEALFMYSKLISGVTEFLNTEGELKELKNFMKNYDGVAAASFRAVETVEANVRWKM  
 LYQDELFQWL GKALRH

**Transmembrane domain:**

amino acids 44-63

**N-glycosylation sites.**

amino acids 89-93, 160-164, 175-179, 222-226, 338-342, 605-609,  
 634-638, 649-653, 663-667, 684-688, 800-804, 906-910

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 362-366

**Tyrosine kinase phosphorylation site.**

amino acids 520-528

**N-myristoylation sites.**

amino acids 78-84, 87-93, 90-96, 118-124, 501-507, 604-610,  
 825-831, 987-993

**Neutral zinc metallopeptidases, zinc-binding region signature.**

amino acids 437-447



**FIGURE 200**

MPAARPPAAGLRGISLFLALLLGSPAAALERDALPEGDASPLGPYLLPSGAPERGSPGKEHPE  
 ERVVTAPPSSSQSAEVLGELVLDGTAPSAHHDIPALSPLLPEEARPKHALPPKKKLPSLKQVN  
 SARKQLRPKATSAATVQRAGSQPASQGLDLLSSSTEKPGPPGDPDPVASEEASEVPLWLDRK  
 ESAVPTTPAPLQISPFTSQPYVAHTLPQRPEPGEPGDMAQEAPQEDTSPMALMDKGENELTG  
 SASEESQETTTSTIITTTVITTEQAPALCSVSFSNPEGYIDSSDYPLLPLNNFLECTYNVTY  
 TGYGVELQVKSVNLSDGELLSIRGVDGPTLTVLANQTLLVEGQVIRSPTNTISVYFRTFQDDG  
 LGTFQLHYQAFMLSCNFPRRPDSGDVTVMDLHSGGVAHFHCHLGYELQGAKMLTCINASKPHW  
 SSQEPICSAPCGGAVHNATIGRVLSPSYPENTNGSQFCIWTIEAPEGQKLHLHFERLLLHDKD  
 RMTVHSGQTNKSALLYDSLQTESVPEGLLSEGNTIRIEFTSDQARAASTFNIRFEAFEKGHC  
 YEPYIQNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPAII ECINVRDPYWNDTEPLCRAMCGG  
 ELSAVAGVVLSPNWPEPYVEGEDCIWKIHVGEKRIFLDIQFLNLSNSDILTIDYDGEVMPHI  
 LGQYLGNSGPQKLYSSTPDLTIQFHSDPAGLIFGKGQGFIMNYIEVSRNDSCSDLPEIQNGWK  
 TTSHTELVRGARITYQCDPGYDIVGSDTLTCQWDLSSWSSDPPFCEKIMYCTDPGEVDHSTRLI  
 SDPVLLVGTTIQYTCNPGFVLEGSSLLTCYSRETGTPIWTSRLPHCVSEESLACDNPGLPENG  
 YQILYKRLLYLPGESLTFMCYEGFELMGEVTIRCILGQPSHWNGPLPVCKVNQDSFEHALEAEA  
 AAETSLEGGMALAIFIPVLIISLLLGGAYIYITRCRYYSNLRLPLMYSHYPYSQITVETEFDN  
 PIYETGETREYEVSI

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 893-915

**N-glycosylation sites.**

amino acids 311-315, 328-332, 350-354, 435-439, 458-462, 474-478,  
 514-518, 576-580, 618-622, 674-678, 742-746

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 188-192

**N-myristoylation sites.**

amino acids 23-29, 87-93, 146-152, 454-460, 475-481, 575-581,  
 629-635, 695-701, 723-729, 766-772, 877-883, 953-959

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 383-394

**FIGURE 201**

**GATG**GCTACGGCAGGGGGTGGCTCTGGGGCTGACCCGGGAAGTCGGGGTCTCCTTCGCCTTCT  
 GTCTTTCTGCGTCCTACTAGCAGGTTTGTGCAGGGGAACTCAGTGGAGAGGAAGATATATAT  
 CCCCTTAAATAAAACAGCTCCCTGTGTTTCGCCTGCTCAACGCCACTCATCAGATTGGCTGCCA  
 GTCTTCAATTAGTGGAGACACAGGGGTTATCCACGTAGTAGAGAAAGAGGAGGACCTACAGTG  
 GGTATTGACTGATGGCCCCAACCCCCCTTACATGGTTCTGCTGGAGAGCAAGCATTTTTACCAG  
 GGATTTAATGGAGAAGCTGAAAGGGAGAACCAGCCGAATTGCTGGTCTTGCAGTGTCTTGAC  
 CAAGCCCAGTCCCTGCCTCAGGCTTCTCTCCTAGTGTACAGTGCCCCAAATGATGGGTTTGGTGT  
 TTACTCCAATTCCTATGGGCCAGAGTTTGTCTACTGCAGAGAAATACAGTGGAATTCGCTGGG  
 CAATGGTTTGGCTTATGAAGACTTTAGTTTCCCCATCTTTCTTCTTGAAGATGAAAATGAAAC  
 CAAAGTCATCAAGCAGTGCTATCAAGATCACAACCTGAGTCAGAATGGCTCAGCACCAACCTT  
 CCCACTATGTGCCATGCAGCTCTTTTCACACATGCATGCTGTCATCAGCACTGCCACCTGCAT  
 GCGGCGCAGCTCCATCCAAAGCACCTTCAGCATCAACCCAGAAATCGTCTGTGACCCCCTGTC  
 TGATTACAATGTGTGGAGCATGCTAAAGCCTATAAATACAACTGGGACATTAAAGCCTGACGA  
 CAGGGTTGTGGTTGCTGCCACCCGGCTGGATAGTCGTTCTTTTTCTGGAATGTGGCCCCAGG  
 GGCTGAAAGCGCAGTGGCTTCCTTTGTACCCAGCTGGCTGCTGCTGAAGCTTTGCAAAGGC  
 ACCTGATGTGACCACCCTGCCCCGCAATGTCATGTTTGTCTTCTTTCAAGGGGAACTTTTGA  
 CTACATTGGCAGCTCGAGGATGGTCTACGATATGGAGAAGGGCAAGTTTCCCGTGCAGTTAGA  
 GAATGTTGACTCATTTGTGGAGCTGGGACAGGTGGCCTTAAGAACTTCATTAGAGCTTTGGAT  
 GCACACAGATCCTGTTTCTCAGAAAAATGAGTCTGTACGGAACCAGGTGGAGGATCTCCTGGC  
 CACATTGGAGAAGAGTGGTGTGGTGTCCCTGCTGTCATCCTCAGGAGGCCAAATCAGTCCCA  
 GCCTCTCCACCATCTTCCCTGCAGCGATTTCTTCGAGCTCGAAACATCTCTGGCGTTGTTCT  
 GGCTGACCACTCTGGTGCCTTCCATAACAAATATTACCAGAGTATTTACGACACTGCTGAGAA  
 CATTAATGTGAGCTATCCCGAATGGCTGAGCCCTGAAGAGGACCTGAACCTTTGTAACAGACAC  
 TGCCAAGGCCCTGGCAGATGTGGCCACGGTGCTGGGACGTGCTCTGTATGAGCTTGCAGGAGG  
 AACCAACTTCAGCGACACAGTTCAGGCTGATCCCCAAACGGTTACCCGCCTGCTCTATGGGTT  
 CCTGATTAAAGCCAACAACCTCATGGTTCCAGTCTATCCTCAGGCAGGACCTAAGGTCCTACTT  
 GGGTGACGGGCCTCTTCAACATTACATCGCTGTCTCCAGCCCCACCAACACCACCTTATGTTGT  
 ACAGTATGCCTTGGCAAATTTGACTGGCACAGTGGTCAACCTCACCCGAGAGCAGTGCCAGGA  
 TCCAAGTAAAGTCCCAAGTGAAAACAAGGATCTGTATGAGTACTCATGGGTCCAGGGCCCTTT  
 GCATTCTAATGAGACGGACCGACTCCCCCGGTGTGTGCGTTCTACTGCACGATTAGCCAGGGC  
 CTTGTCTCCTGCCTTTGAACTGAGTCAGTGGAGCTCTACTGAATACTCTACATGGACTGAGAG  
 CCGCTGGAAAGATATCCGTGCCCGGATATTTCTCATCGCCAGCAAAGAGCTTGAGTTGATCAC  
 CCTGACAGTGGGCTTCGGCATCCTCATCTTCTCCCTCATCGTCACCTACTGCATCAATGCCAA  
 AGCTGATGTCCTTTTTCATTGCTCCCCGGGAGCCAGGAGCTGTGTCATAC**TGA**GGAGGACCCCA  
 GCTTTTCTTGCCAGNTCAGCAGTTCACCTTCCCTAGAGCATCTGTCCCACTGGGACACAACCACT  
 AATTTGTCACTGGAACCTCCCTGGGCCTGTCTCAGATTGGGATTAACATAAAAGAGTGGAACCT  
 ATCCAAAAGAGACAGGGAGAAATAAATAAATTGCTCCCTTCCCTCCGCTCCCCTTTCCCATCA  
 CCCCTTCCCCATTTCCCTCTTCCCTTCTACTCATGCCAGATTTTGGGATTACAAATAGAAGCT  
 TCTTGCTCCTGTTTAACTCCCTAGTTACCCACCCTAATTTGCCCTTCAGGACCCTTCTACTTT  
 TTCTTCCCTGCCCTGTACCTCTCTGCTCCTCACCCCCACCCCTGTACCCAGCCACCTTCCCT  
 GACTGGGAAGGACATAAAAGGTTTAAATGTCAGGGTCAAACCTACATTGAGCCCCTGAGGACAGG  
 GGCATCTCTGGGCTGAGCCTACTGTCTCCTTCCCACTGTCTTTCTCAGGCCCTCAGATGGC  
 ACATTAGGGTGGGCGTGCTGCGGGTGGGTATCCACCTCCAGCCCACAGTGCTCAGTTGTACT  
 TTTTATTAAGCTGTAATATCTATTTTGTTTTTGTCTTTTTCTTTTATTCTTTTGTAAATAT  
 ATATATAATGAGTTTCATTAAAATAGATTATCCC

**FIGURE 202**

MATAGGGSGADPGSRGLLRLLSFCVLLAGLCRGNSVERKIYIPLNKTAPCVRLLNATHQIGCQ  
 SSISGDTGVIHVVEKEEDLQWVLTDGPNPPYMVLLLESKHFTTRDLMEKLGRTSRIAGLAVSLT  
 KPSPASGFSPSVQCPNDGFGVYSNSYGPEFAHCREIQWNSLGNGLAYEDFSFPIFLLLEDENET  
 KVIKQCYQDHNLSQNGSAPTFPLCAMQLFSHMHAVISTATCMRRSSIQSTFSINPEIVCDPLS  
 DYNVWSMLKPINTTGTLPDDRVRVVAATRLDSRSFFWNVAPGAESAVASFVTQLAAAEALQKA  
 PDVTTLPRNVMFVFFQGETFDYIGSSRMVYDMEKGKFPVQLENVDSFVELGQVALRTSLELWM  
 HTDPVSQKNESVRNQVEDLLATLEKSGAGVPAVILRRPNQSQPLPPSSLQRFLRARNISGVVL  
 ADHSGAFHNKYYQSIYDTAENINVSYPEWLSPEEDLNFTVDTAKALADVATVLGRALYELAGG  
 TNFSDTVQADPQTVTRLLYGFLIKANNSWFQSILRQDLRSYLGDGPLQHYIAVSSPTNTTYVV  
 QYALANLTGTVVNLTREQCQDPSKVPSENKDLYEYSWVGPLHSNETDRLPRCVRSTARLARA  
 LSPAFELSQWSSTEYSTWTESRWKDIRARIFLIASKELELITLTVGFGILIEFLIVTYCINAK  
 ADVLFIAPREPGAVSY

**Signal peptide:**

amino acids 1-33

**Transmembrane domain:**

amino acids 671-692

**N-glycosylation sites.**

amino acids 45-49, 55-59, 187-191, 200-204, 204-208, 264-268,  
 387-391, 417-421, 435-439, 464-468, 506-510, 530-534, 562-566,  
 573-577, 580-584, 612-616

**Glycosaminoglycan attachment site.**

amino acids 404-408

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 232-236

**N-myristoylation site.**

amino acids 5-11, 6-12, 9-15, 29-35, 61-67, 120-126, 146-152,  
 168-174, 205-211, 294-300, 438-444, 446-452, 504-510, 576-582



**FIGURE 204**

MRSVQIFLSQCRLLLLLLVPTMLLKSLGEDVIFHPEGEFDSYEVTIPEKLSFRGEVQGVVSPVS  
 YLLQLKGKKHVLHLWPKRLLLPRHLRVFSFTEHGELLEDPYIPKDCNYMGSVKESLDSKATI  
 STCMGGLRGVFNIDAKHYQIEPLKASPSFEHVYLLKKEQFGNQVCGLSDDEIEWQMAPYENK  
 ARLRDFPGSYKHPKYLELILLFDQSRVRFVNNNLSQVIHDAILLTGIMDTYFQDVRMRIHLKA  
 LEVWTDNFKIRVGYPELAEVLGRFVIYKKSVLNARLSSDWAHLYLQRKYNDALAWSFGKVCSL  
 EYAGSVSTLLDTNILAPATWSAHELGHAVGMSHDEQYCQCRGRLNCIMGSGRTGFSNCSYISF  
 FKHISSGATCLNNIPGLGYVLKRCGNKIVEDNEECDGCGSTEECQKDRCCQSNCKLQPGANCSI  
 GLCCHDCRFRPSGYVCRQEGNECDLAEYCDGNSSSCPNDVYKQDGTTPCKYEGRCFRKGCRSRY  
 MQCQSIFGPDAMEAPSECYDAVNLIQDQFGNCEITGIRNFKKCESANSICGRLQCINVTIPD  
 LPEHTTIISTHLQAENLMCWGTGYHLSMKPMGIPDLGMINDGTSCGEGRVCFKKNVCNVSSVLQ  
 FDCLPEKCNTRGVCNNRKNCHCMYGWAPPFCEEVGYGGSIDSGPPGLLRGAIPSSIWVVSIIIM  
 FRLILLILSVVFVFFRQVIGNHLKPKQEKMPLSKAKTEQEESKTKTVQEESKTKTGQEESEAK  
 TGQEESKAKTGQEESKANIESKRPKAKSVKKQKK

**Signal peptide:**

amino acids 1-27

**Transmembrane domain:**

amino acids 684-705

**N-glycosylation sites.**

amino acids 222-226, 372-376, 438-442, 473-477, 625-629

**N-myristoylation sites.**

amino acids 131-137, 168-174, 235-241, 319-325, 364-370, 436-442,  
 472-478, 609-615, 642-648, 668-674, 676-680, 680-686, 749-755,  
 758-764, 767-773

**Amidation site.**

amino acids 69-73

**Disintegrins proteins**

amino acids 429-479

**EGF-like domain proteins**

amino acids 650-662

**Neutral zinc metalloproteinases, zinc-binding region proteins**

amino acids 335-345



FIGURE 205

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGGAAGGTTGAATGGGGTAGAAGGCCTG  
TTGTGGAGGGAAACCACCCATCCTCCTGCCTCCCACCACCACCATCATCCTGGCTGGACGGAG  
AGGGTGACGGGGGCTGGGAAGGGGCAGCTCATGTTACGGTTTCCAGGAGGGGCTACCTGTTGA  
CTGTCTTTGCAGGAAGAAGAAAACACCTGAGTGACCAGATGTCCCAGCTCCAGGTGCCTTGCC  
AGATGGCCAGAACCACACCTCTTGAAGAGTGACAGTGCTGTGGAGCATGGTTTCTGCACACCT  
GGAATGACTGGAACCCCAAAGACTCAAGAAGGAGCTAAAGATCTTGAAGTAGACATGAATAAA  
ACAGAAGGCTGTGGACCACCTGTGAGATGGAGAAGTCTTCTGAGGCTATCCAAACACGGAC  
CAGGCCATGAGACCCCGATGACCATCCCTGAATTTTTTTCGAGAGTCAGTCAACCGATTTGGAA  
CTTATCCAGCCCTCCCATCCAAGAATGGCAAAAAGTGGGAAATTCTGAATTTCAACCAGTACT  
ATGAGGCTTGTCGGAAGGCTGCAAAATCCTTGATCAAGCTGGGTTTGGAGCGTTTCCACGGAG  
TTGGTATCCTGGGGTTTAACTCTGCAGAGTGGTTTATCACTGCTGTTGGTGCCATCCTAGCCG  
GGGGTCTTTGTGTTGGTATTTATGCCACCAACTCTGCCGAGGCTTGTCATATGTCACTACTC  
ATGCCAAAGTGAACATCTTGCTGGTTGAGAATGATCAACAGTTACAGAAAATCCTTTCGATTTC  
CACAGAGCAGCCTAGAGCCCCTAAAAGCGATCATCCAGTACAGACTGCCAATGAAGAAGAACA  
ACAACCTTGTA CTCTTG GGATGATTTT CATGGA ACTTGGCAGAAGTATCCCTGACACCCAACTGG  
AGCAGGT CATCGAGAGCCAGAAGGCGAATCAATGCGCAGTGCTCATCTACACTTCAGGGACCA  
CAGGCATACCCAAAGGGAGTGATGCTCAGTCATGACAACATCACGTGGATTGCAGGAGCAGTGA  
CAAAGGACTTTTAAACTGACAGACAAGCATGAGACGGTGGTTAGCTACCTCCACTCAGCCATA  
TTGCAGCACAGATGATGGACATCTGGGTACCCATAAAGATTGGGGCGCTCACATACTTTGCTC  
AAGCAGATGCTCTCAAGGGCACCTTGTTAAGTACTCTAAAGGAGGTAACCTACTGTCTTCA  
TTGGAGTGCCTCAAATTTGGGAGAAGATACATGAGATGGTGAAGAAAAATAGTGCCAAGTCCA  
TGGGCTTGAAGAAGAAGGCATTTCGTGTGGGCAAGAAACATTGGCTTCAAGGTCAACTCAAAA  
AGATGTTGGGGAAATATAATACTCCCGTGAGCTACCGCATGGCTAAGACTCTCGTGTTCAGCA  
AAGTCAAGACATCCCTTGCTTGATCACTGTCACTCTTTTATCAGTGGGACTGCGCCCCCTCA  
ACCAAGAGACTGCCGAGTTCTTTCTAAGCTTGGACATACCTATAGGCGAGTTGTATGGGTTGA  
GTGAGAGCTCGGGACCCACACGATATCCAACCAGAATAACTACAGGCTTCTAAGCTGTGGCA  
AGATCTTGACTGGGTGTAAGAATATGCTGTTCCAGCAGAACAAGGATGGCATTGGGGAGATCT  
GCCTCTGGGGTAGGCACATCTTCATGGGCTATCTGGAAAGTGAGACTGAAACTACAGAGGCCA  
TCGATGATGAAGGCTGGCTACACTCTGGGGATCTGGGCCAGCTGGACGGTCTGGGTTTCCTCT  
ATGTCACCGGCCACATCAAAGAAATCCTTATCACTGCTGGTGGTGAAAATGTGCCCCCATTC  
CTGTTGAGACCTTGTTAAGAAGAAGATCCCCATCATCAGTAACGCCATGTTAGTAGGAGATA  
AACTGAAGTTTCTGAGCATGTTGCTGACGCTGAAGTGTGAGATGAATCAGATGAGCGGAGAAC  
CTCTGGACAAGCTGAACTTCGAGGCCATCAACTTCTGTGCGGGTCTGGGCAGCCAGGCATCCA  
CCGTGACTGAGATTGTGAAGCAGCAAGACCCCCCTGGTCTACAAGGCCATCCAGCAAGGCATCA  
ATGCTGTGAACCAGGAAGCCATGAACAATGCACAGAGGATTGAAAAGTGGGTCACTTTGGAGA  
AGGACTTTTCCATCTATGGTGGAGAGCTAGGTCCAATGATGAAACTTAAGAGACATTTTGTAG  
CCCAGAAATACAAAAACAAATTGATCACATGTACCCTGACTGCTTTGATGGAGCTGCTCTC  
AGCTGTTCTGATGCCTTCAGCAGGAAGACCTCATTTGCAATAAGTGAAATGCTGCTCTAGGTAG  
AAGCTCTCCCTGCTGTTTTTTAAGAAGCCACATTCTCATTTGGTCAGTTTCTTGATTGTTCTGTC  
TGTTGGAGAGGTGCTCCCTAGAAGAACCTGCCATACGTTTCAAAGCAATAAAATCACTGTATA  
TCTTTCTAAGGACCTTCAAGTCATGACTCCAGGGAAGCCTATTGGGAAGTCTACTAAAACTG  
CCTGATTTTACAAGAAAAGACCTGAACTTGTGGGCTCCCATTGATTTTTTCTCCTCAGGGGAC  
TCAGACATTAGAAAGAAAAAGCCTCACAGATTTGAAGAACTGGACCCCCAAATCAACTCACCT  
GCCTGGAAGCAACTGGGAAACCCTTCCAATAAGTCTTGATAATAAAGCACTTCAGGGTCCCCA  
AAAAAAAAAA

**FIGURE 206**

MTIPEFFRESVNRFGTYPALPSKNGKKWEILNFNQYYEACRKAASLIKLGLERFHGVGILGF  
 NSAWEFITAVGAILAGGLCVGIYATNSAEACQYVITHAKVNILLVENDQQQLQKILSIPQSSLE  
 PLKAI IQYRLPMKKNNNLYSWDDFEMELGRSIPDTQLEQVIESQKANQCAVLIYTS GTTGIPKG  
 VMLSHDNITWIAGAVTKDFKLTDKHETVVSYLPLSHIAAQMMDIWVPIKIGALTYFAQADALK  
 GTLVSTLKEVKPTVFIGVPQIWEKIHVMKKNSAKSMGLKKKAFVWARNIGFKVNSKKMLGKY  
 NTPVSYRMAKTLVFSKVKTSGLGDHCHSFISGTAPLNQETAEFFLSLDIPIGELYGLSESSGP  
 HTISNQNNYRLLSCGKILTGCKNMLFQQNKDGIGEICLWGRHIFMGYLESETETTEAIDDEGW  
 LHSGDLGQLDGLGFLYVTGHIKEILITAGGENVPPIPVETLVKKKIPIISNAMLVGDCLKFSL  
 MLLTLKCEMNQMSGEPLDKLNFEAINFCRGLGSQASTVTEIVKQQDPLVYKAIQQGINAVNQE  
 AMNNAQRIEKWVILEKDFSIYGGELGPMMKLKRHFVAQKYKKQIDHMYH

**Signal peptide:**

amino acids 1-22

**Transmembrane domain:**

amino acids 65-86

**N-glycosylation site.**

amino acids 196-200

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 282-286

**Tyrosine kinase phosphorylation sites.**

amino acids 547-555, 608-616

**N-myristoylation sites.**

amino acids 15-21, 74-80, 80-86, 84-90, 185-191, 189-195,  
 253-259, 337-343, 371-377, 448-454, 536-542

**Amidation site.**

amino acids 24-28

**Putative AMP-binding domain signature.**

amino acids 177-189

**Putative AMP-binding domain proteins.**

amino acids 173-190



**FIGURE 208**

MAYRVLGRAGPPQPRRARRLLFAFTLSLSCTYLCYSFLCCDDLGRSRLLGAPRCLRGPSAGG  
QKLLQKSRPCDPSGPTPSEPSAPSAPAAVPAPRLSGSNHSGSPKLGTKRLPQALIVGVKKGG  
TRAVLEFIRVHPDVRLGTEPHFFDRNYGRGLDWYRSLMPRTLQSITLEKTPSYFVTQEAPR  
RIFNMSRDTKLIVVVRNPVTRAISDYTQTLSSKKPDIPTFEGLSFRNRTLGLVDVSWNAIRIGM  
YVLHLESWLQYFPLAQIHVSGERLITDPAGEMGRVQDFLGIKRFITDKHFYFNKTKGFPCLK  
KTESSLLPRCLGKSKGRTHVQIDPEVIDQLREFYRPNYIKFYETVGQDFRWE

**Signal peptide:**

amino acids 1-33

**N-glycosylation sites.**

amino acids 102-106, 193-197, 235-239, 306-310

**Tyrosine kinase phosphorylation site.**

amino acids 296-305

**N-myristoylation sites.**

amino acids 51-57, 100-106, 121-127, 125-131

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 20-31



**FIGURE 210**

MTCPDKPGQLINWFICSLCVPRVRKLWSSRRPRTRRNLLLGTACAIYLGFLVSQVGRASLQHG  
QAAEKGPHRSDTAEPSFPEIPLDGTLAPPESQGNSTLQPNVVYITLRSKRSKPANIRGTVK  
PKRRKKHAVASAAPGQEALVGPSLQPQEAAREADAVAPGYAQGANLVKIGERPWRLVRGPGVR  
AGGPDFLQPSSRESNIRIYSESAPSWLSKDDIRRMRL LADSAVAGLRPVSSRSGARLLVLEGG  
APGAVLRCGPSPCGLLKQPLDMSEVF AFHLDRI LGLNRTLPSVSRKAEFIQDGRPCPIILWDA  
SLSSASNDTHSSVKLTWGTYQQLLKQKCWQNGRVKPKPESGCTEIHHEWSKMALFDLLQIYN  
RLDTNCCGFRPRKEDACVQNGLRPKCDDQGSAA LAHIIQRKHDPRHLVFIDNKGFFDRSEDNL  
NEKLLEGIKEFPASAVSVLKSQHLRQKLLQSLFLDKVYWESQGGRQGIEKLIDVIEHRAKILI  
TYINAHGVKVLPMNE

**Transmembrane domain:**

amino acids 40-56

**N-glycosylation sites.**

amino acids 98-102, 289-293, 322-326

**N-myristoylation sites.**

amino acids 8-14, 41-47, 97-103, 187-193, 251-257, 252-258,  
287-293, 484-490

**FIGURE 211**

GTGGGGTGGTGAGCGCAGCGCCGAGG**ATG**AGGAGGTGCAACAGCGGCTCCGGGGCCGCGCCGCTCGCTGCTGCTGC  
TGCTGCTGTGGCTGCTCGCGGTTCCCGGCGCTAACGCGGGCCCGCGGTCCGGCGCTCTATTGCGCTTCCGACCCGC  
TGACGCTGCTGCAGGCGGACACGGTGCGCGGCGCGGTGCTGGGCTCCCGCAGCGCCTGGGCCGTGGAGTTCTTCG  
CCTCCTGGTGCGGCCACTGCATCGCCTTCCGCCCCGACGTGGAAGGCGCTGGCCGAAGACGTCAAAGCCTGGAGGC  
CGGCCCTGTATCTCGCCGCCCTGGACTGTGCTGAGGAGACCAACAGTGCAGTCTGCAGAGACTTCAACATCCCTG  
GCTTCCCGACTGTGAGGTTCTTCAAGGCCTTTACCAAGAACGGCTCGGGAGCAGTATTTCCAGTGGCTGGTGCTG  
ACGTGCAGACGCTGCGGGAGAGGCTCATTGACGCCCTGGAGTCCCATCATGACACGTGGCCCCCAGCCTGTCCCC  
CACTGGAGCCTGCCAAGCTGGAGGAGATTGATGGATTCTTTGCGAGAAATAACGAAGAGTACCTGGCTCTGATCT  
TTGAAAAGGGAGGCTCCTACCTGGGTAGAGAGGTGGCTCTGGACCTGTCCCAGCACAAAGGCGTGGCGGTGCGCA  
GGGTGCTGAACACAGAGGCCAATGTGGTGAGAAAGTTTGGTGTACCGACTTCCCCTCTTGCTACCTGCTGTTCC  
GGAATGGCTCTGTCTCCCGAGTCCCGTGCTCATGGAATCCAGGTCTTCTATACCGCTTACCTGCAGAGACTCT  
CTGGGCTCACCAGGGAGGCTGCCCAGACCACAGTTGCACCAACCACTGCTAACAAGATAGCTCCCCTGTTTGGGA  
AATTGGCAGATCGCTCCAAGATCTACATGGCTGACCTGGAATCTGCACTGCACTACATCCTGCGGATAGAAGTGG  
CGAGGTTCCCGGTCTGGAAGGGCAGCGCCTGGTGGCCCTGAAAAAGTTTGTGGCAGTGTGGCCAAGTATTTCC  
CTGGCCGGCCCTTAGTCCAGAACTTCTGCACTCCGTGAATGAATGGCTCAAGAGGCAGAAGAGAAATAAAATTC  
CCTACAGTTTCTTTAAACTGCCCTGGACGACAGGAAAGAGGGTGGCGTTCTTGCCAAGAAGGTGAAGTGGATTG  
GCTGCCAGGGGAGTGAGCCGCATTTCCGGGGCTTTCCCTGCTCCCTGTGGGTCTCTTCCACTTCTTGACTGTGC  
AGGCAGCTCGGCAAAATGTAGACCACTCACAGGAAGCAGCCAAGGCCAAGGAGGTCTCCAGCCATCCGAGGCT  
ACGTGCACACTTCTTCCGCTGCCGAGACTGCGCTAGCCACTTCGAGCAGATGGCTGCTGCCTCCATGCACCGGG  
TGGGGAGTCCCAACGCGCGTGTCTCTGGCTCTGGTCTAGCCACAACAGGGTCAATGCTCGCCTTGCAAGTGGCC  
CCAGCGAGGACCCCAAGTTCCCAAGGTGCAGTGGCCACCCCGTGAACCTTTGTTCTGCCTGCCACAATGAACGCC  
TGGATGTGCCCGTGTGGGACGTGGAAGCCACCCTCAACTTCTCAAGGCCACTTCTCCCAAGCAACATCATCC  
TGGACTTCCCTGCAGCTGGGTGAGTGCCTGCGGAGGGATGTGCAGAATGTGGCAGCCGCCCCAGAGCTGGCGATGG  
GAGCCCTGGAGCTGGAAGCCGGAATTCAACTCTGGACCCTGGGAAGCCTGAGATGATGAAGTCCCCACAAACA  
CACCCCAATGTGCCGGCTGAGGGACCTGAGGCAAGTCGACCCCGAAGCTGCACCCCTGGCCTCAGAGCTGCAC  
CAGGCCAGGACCTCTTGAGCACATGGCAGAGCTTCAGAGGAATGAGCAGGAGCAGCCGCTTGGGCAGTGGCACT  
TGAGCAAGCGAGACACAGGGGCTGCATTGCTGGCTGAGTCCAGGGCTGAGAAGAACCCTCTGGGGCCCTTTGG  
AGGTCAGGCGCGTGGGCCGAGCTCCAAGCAGCTGGTGCACATCCCTGAGGGCCAGCTGGAGGCCCCGAGCTGGAC  
GGGGCCGAGGCCAGTGGCTGCAGGTGCTGGGAGGGGGCTTCTCTTACCTGGACATCAGCCTCTGTGTGGGGCTCT  
ATTCCCTGTCTTTCATGGGCCCTGCTGGCCATGTACACCTACTTCCAGGCCAAGATAAGGGCCCTGAAGGGCCATG  
CTGGCCACCCTGCAGCT**TGA**ACCACCTGGGGAGGAGGCGGGAGAGGGAGCTGCCATCTCTAGGCACCTCAAGCCC  
CCTGACCCCATTCCTTCCCTCCCAACCCCTTGCTCCTTGCTGGCCTAGAAGTGTGGGAAATTCAAGGAAAACGAG  
TTGCTCCAGTGAAGCTTCTTGGGGTTGCTAGGACAGAGAGCTCCTTTGACACAAAAGACAGGAGCAGGGTCCAGG  
TTCCCTGCTGTGCAGGAGGGCAGCCCCGGGCAGTGGGCATAGGGCAGCTCAGTCCCTGGCCTCTTAGCACCAC  
ATTCTGTTTTTTCAGCTTATTTGAAGTCTGCTCATCTCACTGGAGCCTCAGTCTCTCCTGCTTGGTCTTGGC  
CCTCAACTGGGGCAAGTGAAGCCAGAGGAGGGTCCCCCAGCTGGGTGGGCTGGAATGGAATCCTCACTAGCTGC  
TGGGGCTCCGCCCACCCTGCTCCCTTCCGACAATGAAGAAGCCTTTGCACCCCTGGGAGGAAGGACCACCCCGGG  
CCCTCTATGCCTGGCCAGCCTCCAGCTCCTCAGACCTCCTGGGTGGGGTTTGGCTTCAGGGTGGGGTTTGAAGC  
TTCTGGAAGTCGTGCTGGTCTCCAGGTGAGGCAAGCCATGGTTGCTGGGCTGTAGGGTGAGTGGCTTGCTTGGT  
GGGACCTGACGAGTTGGTGGCATGGGAAGGATGTGGGTCTCTAGTGCCTTGCCCTGGCTTAGCTGCAGGAGAAGA  
TGGCTGCTTTCACTTCCCCCATTTGAGCTCTGCTCCCTCTGAGCCTGGTCTTTTGTCTTTTATTTTGGTCTC  
CAAGATGAATGCTCATCTTTGGAGGGTGCCAGGTAGAAGCTAGGGAGGGGAGTGTCTTCTCTCTCAGGTTTTCAC  
CTTCCAGTGTGCAGAAGTTAGAAGGGTCTGGCGGGGCGAGTGCCTTACACATGCTTGATTCCACGCTACCCCT  
GCCTTGGGAGGTGTGTGAATAAATTATTTTGTTAAGGCA

**FIGURE 212**

MRRCNNSGSGPPPSLLLLLLLWLLAVPGANAAPRSALYSPSDPLTLLQADTVRGAVLGSRSAWAV  
EFFASWCGHCIAFAPTWKALAEDVKAWRPALYLAALDCAEETNSAVCRDFNIPGFPTVRFFKA  
FTKNGSGAVFPVAGADVQTLRERLIDALESHHDTWPPACPPLEPAKLEEIDGFFARNNEEYLA  
LIFEKGGSYLGREVALDLSQHKGAVARRVLNTEANVVRKFGVTDFFPCSYLLFRNGSVSRVPVL  
MESRSFYTAYLQRLSGLTREAAQTTVAPTTANKIAPT VWKLADRSKIYMADLESALHYILRIE  
VGRFPVLEGQRLVALKKFVAVLAKYFPGRPLVQNFLHSVNEWLKRQKRNKIPYSFFKTALDDR  
KEGAVLAKKVNWIGCQGSEPHFRGFPCSLWVLFHFLT VQAARQNV DHSQEAAKAKEVLPAIRG  
YVHYFFGCRDCASHFEQMAAASMRVGS PNAAVLWLWSSSHNRVNARLAGAPSEDPQFPKVQWP  
PRELCSACHNERLDVPVWDVEATLNFLKAHFSPSNIILDFPAAGSAARRDVQNVAAPELAMG  
ALELESRNSTLDPGKPEMMKSPTNTTPHVPAEGPEASRP PKLHPGLRAAPGQEPPEHMAELQR  
NEQEQPLGQWHL SKRDTGAALLAESRAEKNRLWGPLEVRRVGRSSKQLVDIPEGQLEARAGRG  
RGQWLQVLGGGFSYLDISLCVGLYLSFMGLLAMYTIFYQAKIRALKGHAGHPAA

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 705-728

N-glycosylation sites.

amino acids 130-134, 243-247, 575-579

Glycosaminoglycan attachment site.

amino acids 6-10

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 644-648

N-myristoylation sites.

amino acids 52-58, 56-62, 196-202, 381-387, 392-398, 448-454,  
468-474, 684-690, 702-708

Cytochrome c family heme-binding site signature.

amino acids 509-515

## Thioredoxin family proteins

amino acids 62-78



**FIGURE 213**

GCACGAGGCCGACTTCCAGACCATCTACAACCTGCACGGCCTGGAACAGCTTCGGCTCCGACAC  
TGAGATCATCCGGCTCAAGGAGCAAGGTTTCGGAATGAAGTCGGGAGCCGGGCTGGAAGCAGA  
GTCTGTGCCGATGGCCGTCATCATTTGGGGTGGCCGTAGGAGCTGGTGTGGCCTTCCTCGTCCT  
TATGGCAACCATCGTGGCGTTCTGCTGTGCCCCGTTCCCAGAGAAATCTCAAAGGTGTTGTGTC  
AGCCAAAAATGATATCCGAGTGGAAATTGTCCACAAGGAACCAGCCTCTGGTCGGGAGGGTGA  
GGAGCACTCCACCATCAAGCAGCTGATGATGGACCGGGGTGAATTCCAGCAAGACTCAGTCCT  
GAAACAGCTGGAGGTCCTCAAAGAAGAGGAGAAAGAGTTTCAGAACCTGAAGGACCCACCAA  
TGGCTACTACAGCGTCAACACCTTCAAAGAGCACCCTCAACCCCGACCATCTCCCTCTCCAG  
CTGCCAGCCCGACCTGCGTCCTGCGGGTAAGCAGCGTGTGCCACAGGCATGTCTTACCAA  
CATCTACAGCACCTGAGCGGCCAGGGCCGCCTCTACGACTACGGGCAGCGGTTTGTGCTGGG  
CATGGGCAGCTCGTCCATCGAGCTTTGTGAGCGGGAGTTCCAGAGAGGCTCCCTCAGCGACAG  
CAGCTCCTTCCTGGACACGCAGTGTGACAGCAGCGTCAGCAGCAGCGGCAAGCAGGATGGCTA  
TGTGCAGTTCGACAAGGCCAGCAAGGCTTCTGCTTCCTCCTCCCACTCCAGTCCTCGTC  
CCAGAACTCTGACCCAGTCGACCCCTGCAGCGGCGGATGCAGACTCACGTCTTAAGGATCACA  
CACCGCGGGTGGGGACGGGCCAGGGAAGAGGTCAGGGCACGTTCTGGTTGTCCAGGGACGAGG  
GGTACTTTGCAGAGGACACCAGAATTGGCCACTTCCAGGACAGCCTCCCAGCGCCTCTGCCAC  
TGCCTTCCTTCGAAGCTCTGATCAAGCACAAATCTGGGTCCCCAGGTGCTGTGTGCCAGAGGT  
GGGCGGGTGGGGAGACAGACAGAGGCTGCGGCTGAGTGCCTGTGCTTAGTGCTGGACACCCG  
TGTCCCCGGCCCTTTCCTGGAGGCCCTCTACCACCTGCTCTGCCCACAGGCACAAGTGGCAG  
CTATAACTCTGCTTTCATGAACTGCGGTCCACTCTCTGGTCTCTCTGTGGGCTCTACCCCTC  
ACTGACCACAAGCTCTACCTACCCCTGTGCCTGTGCTCCCATAACAGCCCTGGGGAGAAGGGGA  
TGACGTCTTCCCAGCACTGAGCTGCCCCAGAAACCCCGGCTCCCCACTGCTGCTCATAGCCCA  
TACCCTGGAGGCTGACAAGCCAGAAATGGCCTTGGCTAAAGGAGCCTCTCTCTACCAGGCTG  
GCCGGGAGCCACCCCAATTTGTTTGGTGTGTTTGTGTCCATACTCTTGCAGTTCTGTCCTTG  
GACTTGATGCCGCTGAACTCTGCGGTGGGACCGGTCCCGTCAGAGCCTGGTGTACTGGGGGA  
GGGAGGGAGGAGGGAGCCTGTGCTGACGGAGCACCTCGCCGGGTGTGCCCCCTCCTGGGCTGTG  
TGACCCAGCCTCCCCACCCACCTCCTGCTTTGTGTACTCCTCCCCTCCCCCTCAGCACAATC  
GGAGTTCATATAAGAAGTGCGGGAGCTTCTCTGGTCAGGGTTCTCTGAACACTTATGGAGAGA  
GTGCTTCCTGGGAAGTGTGGCGTTTGAAGGGGCTGGAGGGCAGGTCTTTAAGATGGCGAGACT  
GCCCTTCTCAGCTGATAAACACAAGAACGGCGATCCTGTCTTCAGTAAGGCTCCACGAGAAGA  
GAGGAAGTATATCTACACCTCAACCTCCTAGTCACCACCTGAAATAAATGTTAGGGAAAAAAA

## **FIGURE 214**

MAVIIGVAVGAGVAFLVLMATIVAFCCARSQRNLKGVVSAKNDIRVEIVHKEPASGREGEEHS  
TIKQLMMDRGEFQQDSVLKQLEVLKEEEKEFQNLKDPTNGYYSVNTFKEHHSTPTISLSSCQP  
DLRPAGKQRVPTGMSFTNIYSTLSGQGRLYDYGQRFVLGMGSSSIELCEREFQRGSLSDSSSF  
LDTQCDSSVSSSGKQDGYVQFDKASKASASSSHHSQSSSQNSDPSRPLQRRMQTHV

**Signal peptide:**

amino acids 1-28

**Glycosaminoglycan attachment site.**

amino acids 150-154

**N-myristoylation sites.**

amino acids 6-12, 10-16, 36-42, 139-145, 165-171

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 114-125

**FIGURE 215**

CAGCCTTCCTCCCCAGCCTGAGTGACTACTCTATTCCCTTGGTCCCTGCTATTGTCGGGGACG  
 ATTGC**ATG**GGCTACGCCAGGAAAGTAGGCTGGGTGACCGCAGGCCTGGTGATTGGGGCTGGCG  
 CCTGCTATTGCATTTATAGACTGACTAGGGGAAGAAAACAGAACAAGGAAAAAATGGCTGAGG  
 GTGGATCTGGGGATGTGGATGATGCTGGGGACTGTTCTGGGGCCAGGTATAATGACTGGTCTG  
 ATGATGATGATGACAGCAATGAGAGCAAGAGTATAGTATGGTACCCACCTTGGGCTCGGATTG  
 GGACTGAAGCTGGAACCAGAGCTAGGGCCAGGGCAAGGGCCAGGGCTACCCGGGCACGTCGGG  
 CTGTCCAGAAACGGGGCTTCCCCCAATTCAGATGATACCGTTTTGTCCCCTCAAGAGCTACAAA  
 AGGTTCTTTGCTTGTTGAGATGTCTGAAAAGCCTTATATTCTTGAAGCAGCTTTAATTGCTC  
 TGGGTAACAATGCTGCTTATGCATTTAACAGAGATATTATTCTGTGATCTGGGTGGTCTCCCAA  
 TTGTCGCAAAGATTCTCAATACTCGGGATCCCATAGTTAAGGAAAAGGCTTTAATTGTCCTGA  
 ATAACCTGAGTGTGAATGCTGAAAATCAGCGCAGGCTTAAAGTATACATGAATCAAGTGTGTG  
 ATGACACAATCACTTCTCGCTTGAACCTCATCTGTGCAGCTTGCTGGACTGAGATTGCTTACAA  
 ATATGACTGTTACTAATGAGTATCAGCACATGCTTGCTAATTCCATTTCTGACTTTTTTTCGTT  
 TATTTTCAGCGGGAAATGAAGAAACCAAACCTTCAGGTTCTGAAACTCCTTTTGAATTTGGCTG  
 AAAATCCAGCCATGACTAGGGAACTGCTCAGGGCCCAAGTACCATCTTCACTGGGCTCCCTCT  
 TTAATAAGAAGGAGAACAAGAAGTTATTCTTAAACTTCTGGTCATATTTGAGAACATAAATG  
 ATAATTTCAAATGGGAAGAAAATGAACCTACTCAGAATCAATTCGGTGAAGGTTCACTTTTTT  
 TCTTTTTAAAAGAATTTCAAGTGTGTGCTGATAAGGTTCTGGGAATAGAAAGTCACCATGATT  
 TTTTGGTGAAAGTAAAAGTTGGAAAATTCATGGCCAAACTTGCTGAACATATGTTCCCAAAGA  
 GCCAGGAAT**TAA**CACCTTGATTTTGTAATTTAGAAGCAACACACATTGTAAACTATTCATTTTC  
 TCCACCTTGTTTATATGGTAAAGGAATCCTTTCAGCTGCCAGTTTTGAATAATGAATATCATA  
 TTGTATCATCAATGCTGATATTTAACTGAGTTGGTCTTTAGGTTTAAGATGGATAAATGAATA  
 TCACTACTTGTTCTGAAAACATGTTTGTTGCTTTTTATCTCGCTGCCTAGATTGAAATATTTT  
 GCTATTTCTTCTGCATAAGTGACAGTGAACCAATTCATCATGAGTAAGCTCCCTTCTGTCATT  
 TTCATTGATTTAATTTGTGTATCATCAATAAAATTGTATGTTAATGCTGGAAAGA

**FIGURE 216**

MGYARKVGWVTAGLVIGAGACYCIYRLTRGRKQNKEKMAEGGSGDVDDAGDCSGARYNDWSDD  
DDDSNESKSIVWYPPWARIGTEAGTRARARARARATRRARRAVQKRASPNSDDTVLSPQELQKV  
LCLVEMSEKPYILEAALIALGNNAAYAFNRDIIRDLGGLPIVAKILNTRDPIVKEKALIVLNN  
LSVNAENQRRLKVYMNQVCDDTITSRLNSSVQLAGLRLLTNMTVTNEYQHMLANSISDFFRLF  
SAGNEETKLQVLKLLLNLAENPAMTRELLRAQVPSSLGSLFNKKENKEVILKLLVIFENINDN  
FKWEENEPTQNQFGEGSLFFFLKEFQVCADKVLGIESHHDFLVKVKVGKFMAKLAEHMFPSQE

**Signal peptide:**

amino acids 1-20

**N-glycosylation sites.**

amino acids 68-72, 189-193, 217-221, 230-234

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 107-111

**N-myristoylation sites.**

amino acids 13-19, 17-23, 19-25, 54-60, 83-89, 147-153, 255-261,  
290-296

**Amidation site.**

amino acids 29-33

FIGURE 217

[illegible]

**FIGURE 218**

MAIAQLATEYVFSDPELLKEPTEPKFKGLRLELAVDKMVTCTIAVGLPLLLISLAFAQEISIGTQ  
ISCFSPSSFSWRQAAFVDSYCWAAVQQKNSLQSESGNLPLWLHKFFPYILLFFAILLYLPPLF  
WRFAAAPHICSDLKFIMEELDKVYNRAIKAASARDLDMRDGACSVPGVTENLGQSLWEVSES  
HFKYPIVEQYLKTKKNSNNLIICYISCRLLTLIIILLACIYLGYYFSLSSLSDEFVCSIKSGI  
LRNDSTVPDQFQCKLIAVGIFQLLSVINLVVYVLLAPVVVYTLFVFPFRQKTDVLKVYEILPTF  
DVLHFKSEGYNDLSLYNLFLEENISEVKSYKCLKVLENIKSSGQGIDPMLLLTNLGMIMDVV  
DGKTPMSAEMREEQGNQTAELQGMNIDSETKANNGEKNARQRLDSSC

**Transmembrane domains:**

amino acids 37-55, 108-126, 216-232, 273-290

**N-glycosylation sites.**

amino acids 255-259, 338-342, 394-398

**Glycosaminoglycan attachment site.**

amino acids 357-361

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 203-207

**N-myristoylation sites.**

amino acids 61-67, 174-180, 251-257, 393-399

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 218-229

**FIGURE 219**

CTGTGAGTGACACACGCTGAGTGGGGTGAAGGGAAATGCTGGTGAATTTTCATTTTGAGGTGTG  
GGTTGCTGTTAGTCACTCTGTCTCTTGCCATTGCCAAGCACAAAGCAATCTTCCTTCACCAAAA  
GTTGTTACCCAAGGGGAACATTGTCCAAGCTGTTGACGCTCTCTATATCAAAGCAGCATGGC  
TCAAAGCAACGATTCCAGAAGACCGCATAAAAAATATACGATTATTAAAAAAGAAAACAAAAA  
AGCAGTTTATGAAAACTGTCAATTTCAAGAACAGCTTCTGTCCTTCTTCATGGAAGACGTTT  
TTGGTCAACTGCAATTGCAAGGCTGCAAGAAAATACGCTTTGTGGAGGACTTTTCATAGCCTTA  
GGCAGAAATTGAGCCACTGTATTTCCCTGTGCTTCATCAGCTAGAGAGATGAAATCCATTACCA  
GGATGAAAAGAATATTTTATAGGATTGGAAACAAAGGAATCTACAAAGCCATCAGTGAAGTGG  
ATATTCTTCTTTCCCTGGATTAAAAAATTATTGGAAAGCAGTCAGTAAACCAAAGCCAAGTACA  
TTGATTTTACAGTTATTTTGAATACAATAAGAACTGCTAGAAATATGTTTATAACAGTCTAT  
TTCTTTTAAAAACTTTTAAACATAATACTGACGGCATGTTAGGTGATTCAGAATAGACAAGAA  
GGATTTAGTAAATTAACGTTTTGGATATAAGTTGTCACTAATTTGCACATTTTCTGTGTTTTT  
AAATAATGTTTCCATTCTGAACATGTTTTGTCATTCACAAGTACATTGTGTCAACTTAATTTA  
AAGTATGTAACCTGAATTAACCTCGTGTAATATTTGTGTGTGGAGTGGGATGTGGGGGGTGGAG  
GGGAATGACAGATTTCTGGAATGCAATGTAATGTTACTGAGACTTAAATAGATGTTATGTAT  
ATGATTGTCTGTTTAAGTGTGTTGAAAATTGTTAATTATGCCCAGTGTGAACTTAGTACTTAAC  
ACATTTTGATTTTAATTAAATAAATTGGGTTTCCTTCTCAAAAAAAAAAAAAAAAAAAAAA  
AAAAA

## **FIGURE 220**

MLVNFILRCGLLLVTLSLAIAKHKQSSFTKSCYPRGTLSQAVDALYIKAAWLKATIPEDRIKN  
IRLLKKKTKKQFMKNCQFQEQLLSFFMEDVFGQLQLQGCKKIRFVEDFHSRLRQKLSHCISCAS  
SAREMKSITRMKRIFYRIGNKGIYKAISELDILLSWIKKLLESSQ

**Signal sequence:**

amino acids 1-21

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 68-71

**N-myristoylation site.**

amino acids 148-153

**Interleukin-10 proteins.**

amino acids 58-94, 74-102, 128-170



**FIGURE 221**

GACCACGGCCCTGCGCCCCAGCCAGGCCTGAGGACATGAGGCGGCCGGCGGGCGGTGCCGCTCC  
TGCTGCTGCTGTGTTTTGGGTCTCAGAGGGCCAAGGCAGCAACAGCCTGTGGTCGCCCCAGGA  
TGCTGAACCGAATGGTGGGCGGGCAGGACACGCAGGAGGGCGAGTGGCCCTGGCAAGTCAGCA  
TCCAGCGCAACGGAAGCCACTTCTGCGGGGGCAGCCTCATCGCGGAGCAGTGGGTCCTGACGG  
CTGCGCACTGCTTCCGCAACACCTCTGAGACGTCCCTGTACCAGGTCCTGCTGGGGGCAAGGC  
AGCTAGTGCAGCCGGGACCACACGCTATGTATGCCCGGGTGAGGCAGGTGGAGAGCAACCCCC  
TGTAACAGGGCACGGCCTCCAGCGCTGACGTGGCCCTGGTGGAGCTGGAGGCACCAGTGCCCT  
TCACCAATTACATCCTCCCCGTGTGCCTGCCTGACCCCTCGGTGATCTTTGAGACGGGCATGA  
ACTGCTGGGTCACTGGCTGGGGCAGCCCCAGTGAGGAAGACCTCCTGCCCGAACCGCGGATCC  
TGCAGAAACTCGCTGTGCCCATCATCGACACACCCAAGTGCAACCTGCTCTACAGCAAAGACA  
CCGAGTTTGGCTACCAACCCAAAACCATCAAGAATGACATGCTGTGCGCCGGCTTCGAGGAGG  
GCAAGAAGGATGCCTGCAAGGGCGACTCGGGCGGCCCCCTGGTGTGCCTCGTGGGTCAGTCGT  
GGCTGCAGGCGGGGGTGATCAGCTGGGGTGAGGGCTGTGCCC GCCAGAACCGCCCAGGTGTCT  
ACATCCGTGTCACCGCCCCACCACAACCTGGATCCATCGGATCATCCCCAACTGCAGTTCCAGC  
CAGCGAGGTTGGGCGGCCAGAAGTGAGACCCCCGGGGCCAGGAGCCCCCTTGAGCAGAGCTCTG  
CACCCAGCCTGCCCCGCCACACCATCCTGCTGGTCCTCCCAGCGCTGCTGTTGCACCTGTGAG  
CCCCACCAGACTCATTTGTAAATAGCGCTCCTTCCTCCCCTCTCAAATACCCTTATTTTATTT  
ATGTTTCTCCCAATAAAAACCCAGCCTGTGTGCCAGCTGAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 222**

MRRPAAVPLLLLLLFCFSQRAKAATACGRPRMLNRMVGGQDTQEGEWPWQVSIQRNGSHFCGGS  
LIAEQWVLTAAHCFRNTSETSLYQVLLGARQLVQPGPHAMYARVRQVESNPLYQGTASSADVA  
LVELEAPVPFTNYILPVCLPDPSVIFETGMNCWVTGWGSPSEEDLLPEPRILQKLAVPIIDTP  
KCNLLYSKDTEFGYQPKTIKNDMLCAGFEEGKKDACKGDSGGPLVCLVGQSWLQAGVISWGEG  
CARQNRPGVYIRVTAHHNWIHRIIPKLQFQPARLGGQK

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **N-glycosylation sites.**

amino acids 55-58, 79-82

#### **Casein kinase II phosphorylation sites.**

amino acids 121-124, 165-168, 167-170, 248-251

#### **Tyrosine kinase phosphorylation sites.**

amino acids 78-86, 197-203

#### **N-myristoylation sites.**

amino acids 16-21, 37-42, 56-61, 62-67, 118-123

#### **Amidation site.**

amino acids 219-222

#### **Serine proteases, trypsin family, histidine active site.**

amino acids 71-76



**FIGURE 224**

MWTALVLIWIFSLSLSESHAASNDPRNFVPNKMWGLVKRNASVETVDNKTSSEVDVTMAAASPV  
 TLTKGTSAAHLNSMEVTTEDTSRTDVSEPATSGVAADGVTSIAPTAVASSTTAASITTAASSM  
 TVASSAPTTAASSTTVASIAPTTAASSMTAASSTPMTLALPAPTSTSTGRTPSTTATGHPSLS  
 TALAQVPKSSALPRTATLATLATRAQTVATTANTSSPMSTRPSPSKHMPSDTAASPVPMPRPQ  
 AQQPISQVSVDQPVVNTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAREPTASPVPVPHTSPIP  
 EMEAMSPTTQPSMPYEQRAAGPGTSQAPEQVETEATPGTDSTGPTPRSSGGTKMPATDSCQP  
 STQGQYMVVTTEPLTQAVVDKTLVVLLLVLTFLITVLVLFALQAYESYKKKDYTQVDYLIN  
 GMYADSEM

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 396-420

**N-glycosylation sites.**

amino acids 41-44, 49-52, 222-225, 268-271, 271-274

**Casein kinase II phosphorylation sites.**

amino acids 14-17, 51-54, 80-83, 85-88, 280-283, 434-437

**N-myristoylation sites.**

amino acids 68-73, 354-359

**Aldo/keto reductase family putative active site signature.**

amino acids 195-210



**FIGURE 226**

MLLPGRARQPPTPQPVQHPGLRRQVEPPGQLLRLFYCTVLVCSKEISALTDGSGYLTKLLQNH  
TTYACDGDYLNLCPRHSTISVQSAFYGQDYQMCSSQKPASQREDSLTCVAATTFQKVLDECQ  
NQRACHLLVNSRVFGPDLCPGSSKYLVSFKCQPNELKNKTVCEDQELKLHCHESKFLNIYSA  
TYGRRTQERDICSSKAERLPPFDCLSYALQVLSRRCYGKQRCKIIVNNHHFGSPCLPGVKKY  
LTVTYACVPKNILTAIDPAIANLKPSLKQKDGEYGINFDPGSGSKVLRKDGILVSNLAAFAYI  
RAHPERAALLFVSSVCIGLALTLCALVIRESCAKDFRDLQLGREQLVPGSDKVEEDSEDEEEE  
EDPSESDFPGELSGFCRTSYPIYSSIEAAELAERIERREQIIQEIWMNSGLDTS�PRNMGQFY

**Transmembrane domains:**

amino acids 32-49, 322-343

**N-glycosylation sites.**

amino acids 62-66, 165-169

**Tyrosine kinase phosphorylation site.**

amino acids 280-287

**N-myristoylation site.**

amino acids 302-308, 333-339, 428-434

**Amidation site.**

amino acids 191-195

**FIGURE 227**

GGCACGAGGTGGAAGGGCTTTTACAAACAGATTGCTGGCCCCACCCCCAGAATTTCTCATCA  
GGAGTGGGCAAGACCAATCATTTGCATTTCTGACAAGTTCCCAGGAGCTGCAGCTGCTGGCCC  
TGGAACCACACTTTGAGAACCACTGCTTTAGACCAAACACCAAAGGAAGATGCAGCCACCCTC  
CTTTACATGTCACAACGCTCAGGGTCCATGAGTACCTCAGGCTGTCCAGCTGAGCTCCACCTG  
CAGCAGCCGAGATTCCCGACTCGCTCCACCATTGGGGGCTAGGAGTGAAGCGTGTACCC**ATGG**  
TCAGCTCATGGCCAGCCAGGAAAGCCTCTCTGCTGTGCGTCTGTGCAGTTCTTGTTCTTCCCT  
GGAGGACTCTTGGATCGCCTGTGATCTTGGCCAGGAGACCAGGTGCCTGGGTCCCTTCCTGGA  
AGGGGACAAGTTACACACCCCAGCCCCATTTTCCACCAACTTCTACATGCCTTGGGAGAACC  
TTCTACATGTTGGCTGCCCCCTTCCCCTATTTTCAGCAGTGCCCAGTCCTGCTTATAAACCTGA  
GGCCTGCTCCCCATACCTTCCCTGTGCAAGTGCCAGCCGTTATTCCAGGCAGCCCAATGTTGT  
TGAGGCCAGATGGATTCCCTGGAAGCAGCTGGCCCATGGATG**TGA**ATCATCACAGTATTCTAGA  
AACAGAGAAGAGGTCTTAACCTAATGCGCATAGAGAAATTGTTCTCATTGTAAACATACCCCT  
GTCCTTAGCTGATCTAGGTGGAAGCCCAGCTTCATGTGCTAGGGGGCATGATAATGATAATAA  
AGGAATTGTATCTAGGACTAA





**FIGURE 229**

GGGAAGGAGATGCAAGGAAGCCCTCCGGCGCTGCGCTCCGAGGCGGGAGACAGCGTCCCGCTGA  
AAATGTGTGTCTGACATGCAAGCTCAGTGGGGCAGAGACCCGTGGATTGCTGTGCCCTGCCCT  
CCGGACCTGGATC**ATG**AAGGTGTTGGGAAGAAGCTTCTTCTGGGTGCTGTTTCCCGTCCTTCC  
CTGGGCGGTGCAGGCTGTGGAGCACGAGGAGGTGGCGCAGCGTGTGATCAAACCTGCACCGCGG  
GCGAGGGGTGGCTGCCATGCAGAGCCGGCAGTGGGTCCGGGACAGCTGCAGGAAGCTCTCAGG  
GCTTCTCCGCCAGAAGAATGCAGTTCTGAACAACTGAAAACCTGCAATTGGAGCAGTGGAGAA  
AGACGTGGGCCCTGTCTGGATGAAGAGAAACTGTTTCAGGTGCACACGTTTGAAATTTTCCAGAA  
AGAGCTGAATGAAAGTGAAATTCGGTTTTTCCAAGCTGTCTACGGACTGCAGAGAGCCCTGCA  
GGGGGATTACAAAGATGTCGTGAACATGAAGGAGAGCAGCCGGGCAGCGCTGGAGGCCCTGAG  
AGAGGCTGCAATAAAGGAAGAAACAGAATATATGGAACCTTCTGGCAGCAGAAAAACATCAAGT  
TGAAGCCCTTAAAAATATGCAACATCAAAACCAAAGTTTATCCATGCTTGACGAGATTCTTGA  
AGATGTAAGAAAGGCAGCGGATCGTCTGGAGGAAGAGATAGAGGAACATGCTTTTGACGACAA  
TAAATCAGTCAAGGGGGTCAATTTTGAGGCAGTTCTGAGGGTGGAGGAAGAAGAGGCCAATTC  
TAAGCAAAATATAACAAAACGAGAAGTGGAGGATGACTTGGGTCTTAGCATGCTGATTGACTC  
CCAGAACAACCAGTATATTTTGACCAAGCCCAGAGATTCAACCATCCCACGTGCAGATCACCA  
CTTTATAAAGGACATTGTTACCATAGGAATGCTGTCTTGCCTTGTGGCTGGCTATGTACAGC  
CATAGGATTGCCTACAATGTTTGGTTATATTATTTGTGGTGTACTTCTGGGACCTTCAGGACT  
AAATAGTATTAAGTCTATTGTGCAAGTGGAGACATTAGGAGAATTTGGGGTGTTTTTTACTCT  
TTTTCTTGTTGGCTTAGAATTTTCTCCAGAAAAGCTAAGAAAGGTGTGGAAGATTTCTTTACA  
AGGGCCGTGTTACATGACACTGTTAATGATTGCATTTGGCTTGTCTGTGGGGGCATCTCTTGCG  
GATCAAACCCACGCAGAGCGTCTTCATTTCCACGTGTCTGTCTTGTCAAGCACACCCCTCGT  
GTCCAGGTTCTCATGGGCAGTGCTCGGGGTGACAAAGAAGGCGACATTGACTACAGCACCGT  
GCTCCTCGGCATGCTGGTGACGCAGGACGTGCAGCTCGGGCTCTTCATGGCCGTGATGCCGAC  
TCTCATACAGGCGGGCGCCAGTGCATCTTCTAGCATTTGTCTGGAAGTTCTCCGAATCCTGGT  
TTTGATTGGTCAGATTCTTTTTTCTACTAGCGGCGGTTTTTCTTTTATGTCTTGTATAAAGAA  
GTATCTCATTGGACCCTATTATCGGAAGCTGCACATGGAAGCAAGGGGAACAAAGAAATCCT  
GATCTTGGAATATCTGCCCTTTATCTTCTTAATGTTAACGGTCACGGAGCTGCTGGACGTCTC  
CATGGAGCTGGGCTGTTTCTGGCTGGAGCGCTCGTCTCCTCTCAGGGCCCCGTGGTCAACGA  
GGAGATCGCCACCTCCATCGAACCCATCCGCGACTTCCTGGCCATCGTTTTCTTCGCCTCCAT  
AGGGCTCCACGTGTTCCCCACGTTTGTGGCGTACGAGCTCACGGTGTGGTGTTCCTCACCTT  
GTCAGTGGTGGTGATGAAGTTTCTCCTGGCGGCGCTGGTCTGTCTCTCATTCTGCCGAGGAG  
CAGCCAGTACATCAAGTGGATCGTCTCTGCGGGGCTTGCCAGGTGAGCAGTTTTCTTTGT  
CCTGGGGAGCCGGGCGGAAGAGCGGGCGTCATCTCTCGGGAGGTGTACCTCCTTATACTGAG  
TGTGACCACGCTCAGCCTCTTGCTCGCCCCGGTGTGTGGAGAGCTGCAATCACGAGGTGTGT  
GCCCAGACCGGAGAGACGGTCCAGCCT**CTGA**TGGCTCGGAGATGATGGACCGTGGAAGGGAAG  
CGTCTGTGGGGAGTGAGCGCTTAGATGGCCAGCAGCTGCTCCTTCTGGGAAGCTCGCACCTTG  
GCAACAGAACAGCCCTCTAGCAGAGCGTCAGTGCAGTCGTGTTATCCCGGCTTTTACAGAATA  
TTCTTGTCTATTTTGAATTTTCCGGAGTAGTTTATTTGCAGTCTGTTGATTATGTGCAGTA  
GACCCGGGACACTGCGTTTTTACCGATCACCTTGAATGTGGTGCCTGGATGTGCCTTTTTTTTT  
TTTCCCTGAAATTATTATTAATTTTCTATTGTGAGTTCATCAGTTCATAGTTTTTTTTTAGTAA  
GAAGCAAAATTAAAAGGCTTTTAAAAATGTACAACCTTCAGAATTATAATCTGTTAGTCAAATA  
TTTGTTATTAAACATTTCTGTAATATGAAGTTGTAATCCTGGCCGTGAGCTTGAAGCTTACT  
TTTGATTCTTAAAGCCTATGTTTTCTAAAATGAGACAAATACGGATGTCTATTTGCCTTTTAT  
TGTAACCTTTTAAATGAAATAATTTTCATGTCAATTTCTATTAGATATATCACTTAAAATATTTG  
GTTTTAAATCACAGAATATGTATTCTTTAATAAAGATAATTTATGATCATGGTAAAAAAAAA

**FIGURE 230**

MKVLGRSFFWVLPVLPWAVQAVEHEEVAQRVIKLHRGRGVAAMQSRQWVRDSCRKLSGLLRQ  
KNAVLNKLKTAIGAVEKDVGLSDEEKLFQVHTFEIFQKELNESENSVVFQAVYGLQRALQGDYK  
DVVNMKESSRQRLEALREAAIKEETEYMELLAAEKHQVEALKNMQHQNQSLSMLDEILEDVRK  
AADRLEEEIEEHAFDDNKS VKGVNFEAVLRVEEEEANSKQNITKREVEDDLGLSMLIDSQNNQ  
YILTKPRDSTIPRADHHFIKDIVTIGMLSLPCGWLCTAIGLPTMFGYIICGVLLGPSGLNSIK  
SIVQVETLGEFGVFFTLFLVGLFESPEKLRKVKISLQGPCYMTLLMIAFGLLWGHLLRIKPT  
QSVFISTCLSLSSTPLVSRFLMGSARGDKEGDIDYSTVLLGMLVTQDVQLGLFMVMPPTLIQA  
GASASSSIVVEVLRILVLIGQILFSLAAVFLCLVIKKYLIGPYRKLHMESSKGNKEILILGI  
SAFIFLMLTVTELDDVSMELGCFLAGALVSSQGPVVTEEIATSIPIRDFLAIVFFASIGLHV  
FPTFVAYELTVLVFLTL SVVVMKFLLAALVLSLILPRSSQYIKWIVSAGLAQVSEFSFVLGSR  
ARRAGVISREVYLLILSVTTLSLLLAPVLWRAAITRCVPRPERRSSL

**Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 282-304, 322-337, 354-370, 379-395, 445-474, 501-520,  
576-598, 641-660

**N-glycosylation sites.**

amino acids 104-108, 174-178, 206-210, 230-234

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 55-59, 673-677

**Tyrosine kinase phosphorylation site.**

amino acids 407-414

**N-myristoylation sites.**

amino acids 116-122, 327-333, 366-372, 401-407, 419-425, 429-435,  
442-448, 525-531, 530-536

**Cell attachment sequence.**

amino acids 404-407

**FIGURE 231**

GAGAAAAACAACAGGAAGCAGCTTACAAACTCGGTGAACAACTGAGGGAACCAAACCAGAGAC  
GCGCTGAACAGAGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTG  
GTGCAAGGCGCAGAGCCAGCCAGATTTGAGAAGAAGGCAAAAAG**ATG**CTGGGGAGCAGAGCTG  
TAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCC  
CTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGCACATC  
CACTAGTGGGACACATGGATCTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCC  
ATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGC  
AAAGGATCCACCAGGGTCTGATTTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGG  
AGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCC  
AACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCC  
AGCCATGGCAGCGTCTCCTTCTCCGCTTCAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTG  
TAGCCGCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCC**TAA**AGGCAGCAGCTCAAGG  
ATGGCACTCAGATCTCCATGGCCCAGCAAGGCCAAGATAAATCTACCACCCCAGGCACCTGTG  
AGCCAACAGGTTAATTAGTCCATTAATTTTAGTGGGACCTGCATATGTTGAAAATTACCAATA  
CTGACTGACATGTGATGCTGACCTATGATAAGGTTGAGTATTTATTAGATGGGAAGGGAAATT  
TGGGGATTATTTATCCTCCTGGGGACAGTTTGGGGAGGATTATTTATTGTATTTATATTGAAT  
TATGTACTTTTTTCAATAAAGTCTTATTTTTTGTGGCTAAAAAAAAAAAAA

## **FIGURE 232**

MLGSRAVMLLLLLPWTAAQGRAVPGGSSPAWTQCQQLSQKLCTLAWSAHPLVGHMDLREEGDEE  
TTNDVPHIQCGDGCDPQGLRDNSQFCLQRIHQGLIFYEKLLGSDIFTGEPSLLPDSPVGQLHA  
SLLGLSQLLQPEGHHWETQQIPSLSPSQPWQRLLLRFKILRSLQAFVAVAAARVFAHGAATLSP

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **Casein kinase II phosphorylation site.**

amino acids 64-67

#### **N-myristoylation sites.**

amino acids 25-30, 81-86, 122-127



**FIGURE 234**

MWAMESGHLLWALLFMQSLWLPQLTDGATRVYYLGIRDVQWNYAPKGRNVITNQPLDSDIVASS  
 FLKSDKNRIGGTYKKTIYKEYKDDSYTDEVAQPAWLGFGLGPVLQAEVGDVILHLKNFATRPY  
 TIHPHGVFYEKDSEGSLYPDGSSGPLKADDSVPPGGSHIYNWTIPEGHAPTDADPACLTWIIYH  
 SHVDAPRDIATGLIGPLITCKRGALDGNSPQQRQDVDHDFLLFSVVDENLSWHLNENIATYC  
 SDPASVDKEDETFQESNRMHAINGFVFGNLPELNMCAQKRVAWHLFGMGNEIDVHTAFFHGQM  
 LTTRGHHTDVANIFPATFVTAEMVPWEPGTWLI SCQVNSHFRDGMQALYKVKSCSMAPPVDLL  
 TGKVRQYFIEAHEIQWDYGPMGHGSGTGKLNREP GSISDKFFQKSSSRIGGTYWKVRYEAFQD  
 ETFQEKMHLEEDRHLGILGPVIRAEVGDITQVVFYNRASQPF SMQPHGVFYEKDYEGTVYNDG  
 SSYPGLVAKPF EKVTYRWTVP PHAGPTAQDPACLTWMYFSAADPIRDTNSGLVGPLLVCRA  
 LGADGKQKGVDKEFFLLFTVL DENKSWYSNANQAAMLDFRLLSE DIEGFQDSNRMHAINGFL  
 FSNLPRLD MCKGDTVAVHLLGLGTETDVHGV MFQGNTVQLQGM RKGAA MLFPHTFVMAIMQPD  
 NLGTFEIYCQAGSHREAGMRAIYNVSQC PGHQATPRQRYQAARIYYIMAEVEWDYCPDRSWE  
 REWHNQSEKDSYGYIFLSNKDGLLSRYKKAVFREYTDGTFRI PRPTGP EEHLGILGPLIKG  
 EVGDILT VVFKNNASRPYSVHAHGVLESTTVWPLAAEPGEVVTYQWNI PERSGPGPNDSACVS  
 WIYYSAVDPIKDMYSGLVGPLAICQKGILEPHGGRSDMDREFALLFLIFDENKSWYLEENVAT  
 HGSQDPGSINLQDET FLESNKMHAINGKLYANLRGLTMYQGERVAWYMLAMQDQDVLHTIHFH  
 AESFLYRNGENYRADVVDLFPGT FEVVEMVASNPGTWLMHCHVT DHVHAGMETLFTVFSRTEH  
 LSPLTVITKETEKVPPRDIEEGNVKMLGMQIPIKNVEMLASVLVAISVTLLLVLALGGVVWY  
 QHRQRKLRRNRRSILDDSFKLLSFKQ

**Signal peptide:**

amino acids 1-21

**Transmembrane domain:**

amino acids 1109-1130

**N-glycosylation sites.**

amino acids 167-171, 239-243, 591-595, 717-721, 761-765, 832-836,  
 876-880, 934-938

**Glycosaminoglycan attachment site.**

amino acids 871-875

**Tyrosine kinase phosphorylation sites.**

amino acids 82-90, 137-145, 494-502, 513-521

**N-myristoylation sites.**

amino acids 212-218, 313-319, 498-504, 566-572, 672-678, 778-784,  
 843-849

**Multicopper oxidases signature 1.**

amino acids 344-365, 696-717, 1043-1064

**Multicopper oxidases signature 2.**

amino acids 1048-1060



**FIGURE 236**

MENLSLSIEDVQPRSPGRSSLDDSGERDEKLSKSISFTSEISIRVSETESFDGNSSKGGLGKE  
ESQNEKQTKKSLLPTLEKKLTRVPSKSLDLNKNEYLSDLKSSTSDSVDEENVPEKDLHGRLFI  
NRIFHISADRMFELLFTSSRFMQKFASSRNIIDVVSTPWTAEELGGDQLRTMTYTIVLNSPLTG  
KCTAATEKQTLTKESREARFYLVDSVLTVDVPHYDYFYTVNRYCIIRSSKQKCRLRVSTD  
YRKQPWGLVKSLIEKNSWSSLEDYFKQLESDDLIEESVLNQAIEDPGKLTGLRRRRRTFNRTA  
ETVPKLSSQHSSGDVGLGAKGDITGKKKEMENYNVTLIVVMSIFVLLLVLNVTFLKLSKIE  
HAAQSFYRLRLQEEKSLNLASDMVSRAETIQKNKDQAHRLKGVLRDSIVMLEQLKSSLIMLQK  
TFDLLNKNKTGMARES

**Transmembrane domain:**

amino acids 352-371

**N-glycosylation sites.**

amino acids 3-7, 54-58, 312-316, 349-353, 367-371, 449-453

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 81-85, 307-311

**Tyrosine kinase phosphorylation sites.**

amino acids 202-211, 246-254, 341-349

**N-myristoylation site.**

amino acids 259-265

**Amidation site.**

amino acids 339-343



**FIGURE 237**

CAGGGGCTGGAGGGCAGGGGAGGGGATGATGTCATTCTGCTCGGCGCAATCCTGACCCTGCT  
CTGGGCGCCACGGCTCAGGCTGAGGTTCTGCTGCAGCCTGACTTCAATGCTGAAAAGTTCTC  
AGGCCTCTGGTACGTGGTCTCCATGGCATCTGACTGCAGGGTCTTCCTGGGCAAGAAGGACCA  
CCTGTCCATGTCCACCAGGGCCATCAGGCCCACAGAGGAGGGCGGCCTCCACGTCCACATGGA  
GTTCCCGGGGGCGGACGGCTGTAACCAGGTGGATGCCGAGTACCTGAAGGTGGGCTCCGAGGG  
ACACTTCAGAGTCCCGGCCTTGGGCTACCTGGACGTGCGCATCGTGGACACAGACTACAGCTC  
CTTCGCCGTCTTTTACATCTACAAGGAGCTGGAGGGGGCCCTCAGCACCATGGTGCAGCTCTA  
CAGCCGGACCCAGGATGTGAGTCCCCAGGCTCTGAAGTCCTTCCAGGACTTCTACCCGACCCT  
GGGGCTCCCCAAGGACATGATGGTCATGCTGCCCCAGTCAGATGCATGCAACCCTGAGAGCAA  
GGAGGCGCCCTGACACCTCCGGAGCCCCACCCCCGCCCTTCCCAGGTGGAGCCAAAGCAGCAG  
GCGCCTTTGCCCCCTGGAGTCAAGACCCACAGCCCTCGGGGACCACCTGGAGTCTCTCCATCCT  
CCACCCCCCGCCTGTGGGATGCCTTGTGGGACGTCTCTTTCTATTCAATAAACAGATGCTGCA  
GCCTCA

## **FIGURE 238**

MMSFLLGAILTLLWAPTAQAEVLLQPDFNAEKFSGLWYVVSMSDCRVFLGKKDHLSMSTRAI  
RPTEEGGLHVHMEFPGADGCNQVDAEYLKVGSEGHFRVPALGYLDVRIVDTDYSSFAVLYIYK  
ELEGALSTMVQLYSRTQDVSPQALKSFQDFYPTLGLPKDMMVMLPQSDACNPESKEAP

### **Signal peptide:**

amino acids 1-20

### **Tyrosine kinase phosphorylation site.**

amino acids 110-117

### **N-myristoylation sites.**

amino acids 7-13, 79-85, 130-136

### **Amidation site.**

amino acids 50-54

105441-240000

FIGURE 239

GCGCGCGCTGGTCCAGGTGAGCGGGCGCGCTCCCCGCGACGGCGCTGCCTGCCCAGGCGGTTCA  
 CGTAAAGACAGCGAGATCCTGAGGGCCAGCCGGGAAGGAGGCGTGGATATGGAGCTGGCTGCT  
 GCCAAGTCCGGGGCCCCGCGCCGCTGCCTAGCGCGTCTTGGGGACTCTGTGGGGACGCGCCCCG  
 CGCCGCGGCTCGGGGACCCGTAGAGCCCGGCGCTGCGCGC**ATG**GCCCTGCTCTCGCGCCCCGC  
 GCTCACCCCTCCTGCTCCTCCTCATGGCCGCTGTTGTCAGGTGCCAGGAGCAGGCCAGACCAC  
 CGACTGGAGAGCCACCCTGAAGACCATCCGGAACGGCGTTTCATAAGATAGACACGTACCTGAA  
 CGCCGCCTTGGACCTCCTGGGAGGCGAGGACGGTCTCTGCCAGTATAAATGCAGTGACGGATC  
 TAAGCCTTTCCACGTTATGGTTATAAACCTCCCCACCGAATGGATGTGGCTCTCCACTGTT  
 TGGTGTTTCATCTTAACATTGGTATCCCTTCCCTGACAAAGTGTTGCAACCAACACGACAGGTG  
 CTATGAGACCTGTGGCAAAAGCAAGAATGACTGTGATGAAGAATTCCAGTATTGCCTCTCCAA  
 GATCTGCCGAGATGTACAGAAAACACTAGGACTAACTCAGCATGTTTCAGGCATGTGAAACAAC  
 AGTGAGGCTCTTGTTTGACAGTGTTATACATTTAGGTTGTAAACCATATCTGGACAGCCAACG  
 AGCCGCATGCAGGTGTCATTATGAAGAAAAAACTGATCTTT**TAA**AGGAGATGCCGACAGCTAGT  
 GACAGATGAAGATGGAAGAACATAACCTTTGACAAATAACTAATGTTTTTACAACATAAACT  
 GTCTTATTTTTTGTAAGGATTATTTTGAGACCTTAAAATAATTTATATCTTGATGTTAAAC  
 CTCAAAGCAAAAAAAGTGAGGGAGATAGTGAGGGGAGGGGCACGCTTGCTCTCTCAGGTATCTT  
 CCCCAGCATTGCTCCCTTACTTAGTATGCCAAATGTCTTGACCAATATCAAAAACAAGTGCTT  
 GTTTAGCGGAGAATTTTTGAAAAGAGGAATATATAACTCAATTTTCACAACCACATTTACCAAA  
 AAAAGAGATCAAATATAAAATTCATCATAATGTCTGTTCAACATTATCTTATTTGGAAAATGG  
 GGAAATTTATCACTTACAAGTATTTGTTTACTATGAAATTTTAAATACACATTTATGCCTAGAA  
 GGAACGGACTTTTTTTTTCTATTTTAATTACACATAATATGTAATTAAAGTACAACATAATAT  
 GTTGTTTCTCTGTAGCCCGTTGAGCATATGAGTAAGTCACATTTCTATTAGGACTACTTACAA  
 GGACAAGGTTTCCATTTTTTCCAGTTGTAAATTGGAACCATCAGCTGATAACCTCGTAGGGAG  
 CAACCCCAGGATAGCTAAGTGTTATGTAATATGCCTAGAAGGTGATGTGAATGCGATTCAGAA  
 GCATAGCCACTCCCATTTTTATGAGCTACTCACATGACAAATGTCATCTTTTGCTATAACCTTT  
 GCCAAGTTAGAGAAAAGATGGATTTAATGAGATAAATGAAAAGATATTTAACCTAAAAAAAAA  
 AAAAAAAAAAAAAAAAAA

## **FIGURE 240**

MALLSRPALTLLLLLMAAVVRCQEQATTDWRATLKTIRNGVHKIDTYLNAALDLLGGEDGLC  
QYKCDGSKPFPFRYGYKPSPPNGCGSPLFGVHLNIGIPSLTKCCNQHDRCYETCGKSKNDCDE  
EFQYCLSKICRDVQKTLGLTQHVQACETTVELLFDSVIHLGCKPYLDSQRAACRCHYEKTDL

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **N-myristoylation sites:**

amino acids 57-63,93-99

#### **Phospholipase A2 histidine active site:**

amino acids 106-114

#### **Neuraxin and MAP1B proteins repeat proteins Block:**

amino acids 109-137

**FIGURE 241**

GATTCCGAGCGCCTCCACTGCTGGTCCGTTGGCCAGATCAACTCGCCGCGTGGGCCGGCCGTT  
CCCTGAGAGTCTGAGCGCTCGCCGCACCCCTTCCGAGCTTCTATTGGCCGTAGCAGACGTCC  
GTCTGCCGCTATCTCCGCCCCAATACGGAAGCGGCCTAGTCCTCCGGCTCCGACAGCTGGGTG  
TCCAGGCCATGGGGCAGCCCTGGGCGGCTGGGAGCACGGACGGGGCGCCCGCGCAGCTGCCTC  
TCGTGCTCACCGCGCTGTGGGCCGCGGCCGTGGGCCTGGAGCTGGCTTACGTGCTGGTGCTCG  
GTCCCGGGCCGCCGCCGCTGGGACCCCTGGCCCGGGCCTTGCAGCTGGCGCTGGCCGCCTTCC  
AGCTGCTCAACCTGCTGGGCAACGTGGGGCTCTTCCTGCGCTCGGATCCCAGCATCCGTGGCG  
TGATGCTGGCCGGCCGCGGTCTGGGCCAGGGCTGGGCTTACTGCTACCAATGCCAAAGCCAGG  
TGCCGCCACGCAGCGGACACTGCTCTGCCTGCCGCGTCTGCATCCTGCGTCGGGACCACCACT  
GCCGCCTGCTGGGCCGCTGCGTGGGCTTCGGCAACTACCGGCCCTTCCTGTGCCTGCTGCTTC  
ATGCCGCCGGCGTCCTGCTCCACGTCTCTGTGCTGCTGGGCCCTGCACTGTCGGCCCTGCTGC  
GAGCCCACACGCCCCTCCACATGGCTGCCCTCCTCCTGCTTCCCTGGCTCATGTTGCTCACAG  
GCAGAGTGTCTCTGGCACAGTTTGCCTTGGCCTTCGTGACGGACACGTGCGTGGCGGGTGCGC  
TGCTGTGCGGGGCTGGGCTGCTCTTCCATGGGATGCTGCTGCTGCGGGGCCAGACCACATGGG  
AGTGGGCTCGGGGCCAGCACTCCTATGACCTGGGTCCCTGCCACAACCTGCAGGCAGCCCTGG  
GGCCCCGCTGGGCCCTCGTCTGGCTCTGGCCCTTCCTGGCCTCCCCATTGCCTGGGGATGGGA  
TCACCTTCCAGACCACAGCAGATGTGGGACACACAGCCTCCTGACTCCAGGAAGAGCCAGAGC  
TGTGCAGGGAGGAAGGGGTGAGAGGGGGGCCCCACACCTAGACTCAGTAAGGAAGTCGGGTT  
GGACCTTAACATCTGCATTGGACAACCTCCACCCCTTCCTTGGCCTTGCCCCCTGCCCGCCTACA  
CTCCTACGTGTCCAGGGCTTGGGCCGTGACTTAGGCAGAGGAGTGCAGAGGAGGGTCTGGCAG  
GGGCTGCTCAGGCCGCCTAGCTGCCCCCTTGGCAGGTTAATAAAGCACTGACTTGTTAA

**FIGURE 242**

MGQPWAAGSTDGAPAQPLPLVLTALWAAVGLLAYVLVLGPGPPPLGPLARALQLALAAFQLL  
NLLGNVGLFLRSDPSIRGVMLAGRGLGQGWAYCYQCQSQVPPRSGHCSACRVCILRRDHHCLL  
LGRCVGFGNYPFLCLLLHAAGVLLHVSVLLGPALSALLRAHTPLHMAALLLLPWLMLLTGRV  
SLAQFALAFVTDTCVAGALLCGAGLLFHGMLLLRGQTTWEWARGQHSYDLGPCHNLQAALGPR  
WALVWLWPFLASPLPGDGITFQTTADVGHSTAS

**Important features:****Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 51-66,143-160,174-191,198-214

**N-myristoylation sites:**

amino acids 2-8,8-14,30-36,81-87,88-94,90-96,206-212

**Leucine zipper pattern:**

amino acids 143-165,150-172,157-179,164-186

**FIGURE 243**

CTTGTCTTTGTGTCGGTTGTGATTTTCCCTAATCTCTGATTTTCCCTTTCTCTCGGACGCTCTC  
CCTCTTCGGACCCATTTTCTCCCGTGCTTCATGCCCTGATAGCCTGGCCCCCTTCCCGGCTTCC  
TTCGCTACCGGGGACGCCTCTAGTTTTTCTGAATTTCTGGCTGGCTCCACCCTCCGCGTTCAT  
CTTCCTCAAGAGTTTCGCCCCCTCTGGGGGCTCCTCTGTGTAATCGTCGCCTTCTCTGGGTATTT  
CTGTGAACTCCGTCTCACACCATCCCGCCATCTTCTCTGCCTTGGCCCCCTTTTCTCTGTACAG  
CCAGCTCTGTGTCCTTTTCTTCTCCCCCTCTAAAATCGACTCCTCTTCTCCCTGAGAGCCCCA  
CCTTTGTGCCCCACTCCTCATTTTCTACGCCTCCCTCTCTCTGCTGGTCCTCTCTCTCCCTG  
CAAGGTTCCATTCCATCAATTTGTTTGTCTTTTGTAGGGGTGGCATCCCCTCTGACTACTGCT  
CCATCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGCTTGAGGATTTCACTTCAATCTTTTCTGGT  
TGCGTCTCCACTTGTACTCAGCTTGTTAGGTCCAGGTCCAGTTGTTCTGCATCTGAGGCTGGC  
GTGTGCTGTCTTCTCTGATTGGCCTAATCTCCCTCACCCCGTGAGATCTGTTGTCAGCCTTC  
GTTTCTCTTTCTGTGTCCCAGCTTTTCTGCGGGTCTTGGCACCTTTCTTGGCCACAGATTTCT  
TGGGTTACAGAGCATGTGTGTCTGAGGCATTGCAGGCAGAAAAGGTGGCCGACGTGACCTCT  
AGCTGGACTGCTGGGCAGGGGAGCTGTCTAGATAAAAATTGGAAAGAAACAGTGACCCAGAGA  
CAGGTGGACAAAGAATTCGGGGACTGATGGGAACTGAGCTTGGGATCCAGACTGAAACTGATT  
CCAGACTGACCTCTAGCACCCAGGACCCAGACACAGGGGCC**ATGG**GACCCCAGCATTTGAGACT  
TGTGCAGCTGTTCTGCCTTCTAGGGGCCATCCCCACTCTGCCTCGGGCTGGAGCTCTTTTGTG  
CTATGAAGCAACAGCCTCAAGATTCAGAGCTGTTGCTTTCCATAACTGGAAGTGGCTTCTGAT  
GAGGAACATGGTGTGTAAGCTGCAAGAGGGCTGCGAGGAGACGCTAGTGTTCAATTGAGACAGG  
GACTGCAAGGGGAGTTGTGGGCTTTAAAGGCTGCAGCTCGTCTTCGTCTTACCCTGCGCAAAT  
CTCCTACCTTGTTTCCCCACCCGAGTGTCCATTGCCTCCTACAGTCGCGTCTGCCGGTCTTA  
TCTCTGCAACAACCTCACCAATTTGGAGCCTTTTGTGAACTCAAGGCCAGCACTCCTAAGTC  
TATCACATCTGCGTCCTGTAGCTGCCCCACCTGTGTGGGCGAGCACATGAAGGATTGCCTCCC  
AAATTTTGTCAACCACTAATTCTTGCCCCCTTGGCTGCTTCTACGTGTTACAGTTCCACCTTAAA  
ATTTCAAGGCAGGGTTTCTCAATACCACCTTCCTCCTCATGGGGTGTGCTCGTGAACATAACCA  
GCTTTTAGCAGATTTTCATCATATTGGGAGCATCAAAGTGACTGAGGTCCTCAACATCTTAGA  
GAAGTCTCAGATTGTTGGTGCAGCATCCTCCAGGCAAGATCCTGCTTGGGGTGTGCTCTTAGG  
CCTCCTGTTTGCCTTCAGGGAC**TGA**CCATCTAGCTGCACCCGACAAGCACCCAGACTCTTTCA  
CATAACAAATAAAATAGCAGAGTTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAA

## **FIGURE 244**

MGPQHLRLVQLFCLLGAIPTLPRAGALLCYEATASRFRAVAFHNWKWLLMRNMVCKLQEGCEE  
TLVFIETGTARGVVGFKGCSSSSSYPAQISYLVSPPGVSIASYSRVCRSYLCNNLTNLEPFVK  
LKASTPKSITSASCSCPTCVGEHMKDCLPNFVTTNSCPLAASTCYSSTLKFQAGFLNTTFLLM  
GCAREHNQLLADFHIGSIKVTEVLNILEKSQIVGAASSRQDPAWGVVLGLLFAFRD

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **N-glycosylation sites:**

amino acids 117-121,183-187

#### **N-myristoylation sites:**

amino acids 16-22,25-31,60-66,71-77,81-87,100-106,224-230,  
235-241,239-245

#### **Prokaryotic membrane lipoprotein lipid attachment site:**

amino acids 181-192



**FIGURE 245**

GTGGAGTTGGGTGGTGTCTGGGAGCCTCTCCCTGAGGGGCACCGCTCTTCAGGAGCTGGGCCTCCAGTGCGGCGC  
GATGTCAGGCGCGGTGACAGCTCTGTGAGTCCGAGGCCGCGCCGTGGCGCTGGGCGGCTGCGGGGCTGACCGG  
TCCGCTC**ATG**GTGCCGCCACGACGCCATCGCGGGGCAGGAAGGCCAGGGGTGCTGAGTCTTCACCTCCTTTTAG  
ACTGAGATCTGCCAAGTTTTCCGGCATTGCTCTTGAGGATCTCAGAAGGGCTCTTAAGACAAGACTGCAAATGGT  
GTGTGTATTTGTCATGAACGAATGAATTCACAGAACAGTGGTTTTCACTCAGCGCAGGCGAATGGCTCTTGGGAT  
TGTTATTCTTCTGCTTGTTGATGTGATATGGGTTGCTTCTCTGAACCTACTTCGTATGTTTTTACCCAGTACAA  
CAAACCATTCCTTCAGCACCTTTGCAAAAACATCTATGTTTGTGTTTGTACCTTTTGGGCTTTATTATTTGGAAGCC  
ATGGAGACAACAGTGTACAAGAGGACTTCGCGGAAAGCATGCTGCTTTTTTTGTCAGATGCTGAAGGTTACTTTGC  
TGCTTGCACAACAGATACAACATGAATAGTTCTTTGAGTGAACCTCTGTATGTGCCTGTGAAATTCCATGATCT  
TCCAAGTGAAAAACCTGAGAGCACAAACATTGATACTGAAAAAACCCCAAAAAGTCTCGTGTGAGGTTTCAGTAA  
TATCATGGAGATTCGACAGCTTCCGTCAAGTCATGCATTGGAAGCAAAGTTGTCTCGCATGTGCATATCCTGTGAA  
AGAACAAGAATCCATACTGAAAACGTGGGGGAACTTACTGCAACTCAAGTAGCGAAAATTAGCTTTTTTTTTTTG  
CTTTGTGTGGTTTTTTGGCAAATTTGTGCATATCAAGAAGCACTTTCAGACACACAAGTTGCTATAGTTAATATTTT  
ATCTTCAACTTCCGGACTTTTTACCTTAATCCTTGCTGCAGTATTTCCAAGTAACAGTGGAGATAGATTTACCTT  
TTCTAAACTATTAGCTGTAATTTTAAGCATTGGAGGCGTTGTACTGGTAAACCTGGCAGGGTCTGAAAAACCTGC  
TGGAAGAGACACAGTAGGTTCCATTTGGTCTCTTGCTGGAGCCATGCTCTATGCTGTCTATATTGTTATGATTAA  
GAGAAAAGTAGATAGAGAAGACAAGTTGGATATTCCAATGTTCTTTGGTTTTGTAGGTTTGTGTTAATCTGCTGCT  
CTTATGGCCAGGTTTCTTTTTACTTCATTATACTGGATTTGAGGACTTCGAGTTTCCCAATAAAGTAGTATTAAT  
GTGCATTATCATTAATGGCCTTATTGGAACAGTACTCTCAGAGTTCCTGTGGTTGTGGGGCTGCTTTCTTACCTC  
ATCATTGATAGGCACACTTGCACTAAGCCTTACAATACCTCTGTCCATAATAGCTGACATGTGTATGCAAAAGGT  
GCAGTTTTCTTGGTTATTTTTTGCAGGAGCTATCCCTGTATTTTTTTTCATTTTTTTATTGTAACCTCCTATGCCA  
TTATAATAATTGGGATCCTGTGATGGTGGGAATCAGAAGAATATTTGCTTTTATATGCAGAAAACATCGAATTCA  
GAGAGTTCAGAAAGACAGCGAACAGTGTGAGAGTCTCATTTCTATGCACAGTGTCTCAGGAGGATGGAGCTAG  
**T****TAG**CTGTCTGTTGTCTGTAGCCCAGCTTGATAATGGAACATACAGCGAAGAGACAATCTCTGGCAAGTTTTTG  
TAGAAAAAATGTTTCAGTGCCTAGTCTGAAAAATAACAGTTTGAGTTCCTTGAAACTCTAAAATATATTTTTCTC  
ATACCTGTTTTCTTCATTTTCATAATGAAGCACTTTGCTATGTAGCTGTGTACATATCACTACAGTTATAGGAAG  
TTTCAGTCTACAGTCCATCCAAAGGACCAACCTGCCTTACACATCTCAAGGAATTGAGTGTGAAATCATTGGA  
ACTAATCAAGGAATAAATCCTAATGTTCTGGGACTTTATTTTCACATGTTAAATGCTGGAATATATTATGAAAAAT  
GTTTTCAAGAAATCACTTAAGTGTTCATAGACCAGTATTTCTGACAGGTAAAATGCTAAAATAAGCTACCTGTAA  
TAAGTGTGGATTATATTTTTTGGGTTTTGTAGAATATTGCAAAATTAACCACACAAAAAATGTTTAATTTATGCAAC  
AAGCATGTTTGTGCAAATTTTCATGGGACTTTAAAAAGAATAAGTATTTGAGAAAAATATCTGGTTCACCTTACACTA  
CATTTACTGTATTATTCTTTTATAGCATTAGGTGCCTTGTATTTTAAATCTGTGACAAACCATGGCAAATTTTTTA  
AAGGGGAAGTATTATTATAAAATGAAGAAATATGTATTTCTAAAGGCTATATTGCTGTAAACTTAATTGATAAG  
CTCTGTTTAATTTAGAGTTTTGAAGAAATAGTCTCCCTTCAATTAAGAAATTTTCATAATGGAATGATTTAAATT  
GAAGTGACAAAGAGTATTATTAAAATACAATGTTTATAAAAAA

**FIGURE 246**

MVPPRRHRGAGRPGVLSSSPFRLRSKFSGLALEDLRRALKTRLQMVCFVFMNRMNSQNSGF  
TQRRRMALGIVILLLLVDVIWVASSELTSYVFTQYNKPFFSTFAKTSMFVLYLLGFIIWKPWRQ  
QCTRGLRGKHAFFADAEGYFAACTTDTTMNSSLSEPLYVPVKFHDLPSEKPESTNIDTEKTP  
KKSRRVRFNIMEIRQLPSSHALEAKLSRMSYPVKEQESILKTVGKLTATQVAKISFFFCFVWF  
LANLSYQEALSDTQVAIVNILSSTSGLFTLILAAVFPSNSGDRFTLSKLLAVILSIGGVVLVN  
LAGSEKPAGRDTVGSIWLAGAMLYAVYIVMIKRKVDREDKLDIPMFFGFVGLFNLLLLWPGE  
FLLHYTGFEDEFEPNKVVLMCIIINGLIGTVLSEFLWLWGCFLTSSLIGTLALSLTIPLSIIA  
DMCMQKVQFSWLFFAGAIPVFFSFFIVTLLCHYNNWDPVMVGIRRIFAFICRKHRIQRPEDS  
EQCESLISMHSVSQEDGAS

**Important features:****Transmembrane domain:**

amino acids 69-87,105-118,237-256,266-285,300-316,332-346,  
364-379,399-419,453-472

**N-glycosylation sites:**

amino acids 157-161,255-259

**N-myristoylation sites:**

amino acids 14-20,329-335,404-410,407-413,418-424

**FIGURE 247**

CGTCTGTAGAGATATCATGAACTTCAACTTAGCTTTGGTACTTTCTTCCCTGAAGACAGAGGG  
CAGAACTCTGAGTTCCAGAACCATTTTCAACTGTATTGGGGACCAATCACTTGACTCTATTCT  
TGTCTCTCTGACAGATGACGCTACACTCTCCTCTGAATAATGGACACCATTTCTAAACTGAA  
TCCTGCTACTAAAATAATTCAGATGATATATTTTCCAATTCTACAATCTTGCTTTGTTTTAT  
TTAGTTGTTTTCTCTCTCTCTTCCCAGTTTTTCCAGAGACTGGAGCTAAACTGGGCTTTCAACA  
TCATCATGAAGTTTATCCTCCTCTGGGCCCTCTTGAATCTGACTGTTGCTTTGGCCTTTAATC  
CAGATTACACAGTCAGCTCCACTCCCCCTTACTTGGTCTATTTGAAATCTGACTACTTGCCT  
GCGCTGGAGTCCTGATCCACCCGCTTTGGGTGATCACAGCTGCACACTGCAATTTACCAAAGC  
TTCGGGTGATATTGGGGGTACAATCCCAGCAGACTCTAATGAAAAGCATCTGCAAGTGATTG  
GCTATGAGAAGATGATTCATCATCCACACTTCTCAGTCACTTCTATTGATCATGACATCATGC  
TAATCAAGCTGAAAACAGAGGCTGAACTCAATGACTATGTGAAATTAGCCAACCTGCCCTACC  
AAACTATCTCTGAAAATACCATGTGCTCTGTCTCTACCTGGAGCTACAATGTGTGTGATATCT  
ACAAAGAGCCCGATTCACTGCAAACGTGTAACATCTCTGTAATCTCCAAGCCTCAGTGTGCGG  
ATGCCTATAAAACCTACAACATCACGGAAAATATGCTGTGTGTGGGCATTGTGCCAGGAAGGA  
GGCAGCCCTGCAAGGAAGTTTCTGCTGCCCCGGCAATCTGCAATGGGATGCTTCAAGGAATCC  
TGTCTTTTGC GGATGGATGTGTTTTGAGAGCCGATGTTGGCATCTATGCCAAAATTTTTTACT  
ATATACCCTGGATTGAAAATGTAATCCAAAATAACTGAGCTGTGGCAGTTGTGGACCATATGA  
CACAGCTTGTCCCCATCGTTCACCTTTAGAATTAAATATAAATTAACCTCCTC

FIGURE 248

MKFILLWALLNLTVAFNPDYTVSSTPPYLVYLKSDYLP CAGVLIHPLWVITAAHCNLPKLR  
VILGVTIPADSNEKHLQVIGYEKMIHHPFSVTSIDHDIMLIK LKTEAELNDYVKLANLPYQT  
ISENTMCSVSTWSYNVCDIYKEPDSLQTVNISVISKPQCRDAYKTYNITENMLCVGIVPGRRQ  
PCKEVSAA PAICNGMLQGILSFADGCVLRADVGIYAKIFYIYPWIENVIQNN

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation sites:

amino acids 11-15, 156-160, 173-177

**Tyrosine kinase phosphorylation site:**

amino acids 108-117

N-myristoylation sites:

amino acids 182-188, 203-209

Amidation site:

amino acids 185-189

Serine proteases, trypsin family, histidine active site:

amino acids 52-58

**FIGURE 249**

GCGAGGCGGCCGCTGTCTTCTGCTGCGGCTTCCGCGACCACAAGTACTGCTGCGACGACCCGC  
ACAGCTTCTTCCCCTACGAGCACAGCTACATGTGGTGGCTCAGCATTGGCGCTCTCATAGGCC  
TGTC CGTAGCAGCAGTGGTTCTTCTCGCCTTCATTGTTACCGCCTGTGTGCTCTGCTACCTGT  
TCATCAGCTCTAAGCCCCACACAAAGTTGGACCTGGGCTTGAGCTTACAGACAGCAGGCCCTG  
AGGAGGTTTCTCCTGACTGCCAAGGTGTGAACACAGGCATGGCGGCAGAAGTGCCAAAAGTGA  
GCCCTCTCCAGCAGAGTTACTCCTGCTTGAACCCGCAGCTGGAGAGCAATGAGGGGCAGGCTG  
TGA ACTCCAAACGCCTCCTCCATCATTGCTTCATGGCCACAGTGACCACCAGTGACATTCCAG  
GCAGCCCTGAGGAAGCCTCTGTACCCAACCCTGACCTATGTGGACCAGTCCCATAAACATTCA  
ATAAATGTCTCCATACCATCAA

250/550

## **FIGURE 250**

MWWLSIGALIGLSVAAVLLAFIVTACVLCYLFISSKPHTKLDLGLSLQTAGPEEVSPDCQGV  
NTGMAAEVPKVSPLQQSYSCLNPOLESNEGQAVNSKRLLLHCFMATVTTSDIPGSPEEASVPN  
PDL CGPVP

**Important features:**

**Signal peptide:**

Amino acids 1-26

**N-myristoylation sites:**

Amino acids 7-13, 11-17, 62-68, 93-99

250/550

FIGURE 251

GTGGTTTGGATTGAGCCGGGCCGGCCGGGGGCCCGAGTCGGAGGGGGGTGGCAGTGAGCGGCC  
GCAGAGGCTACGGGGCTCGGTTTGGCTGACTGGGGAGTCGGCAGGCGGCAGGAACC**ATG**CGAG  
GCCAGCGGAGCCTGCTGCTGGGCCCCGGCCCCGCTCTGCCTCCGCCTCCTTCTGCTGCTGGGTT  
ACAGGCGCCGCTGTCCACCTCTACTCCGGGGTCTAGTACAGCGCTGGCGCTACGGCAAGGTCT  
GCCTGCGCTCCCTGCTCTACAACTCCTTTGGGGGCAGTGACACCGCTGTTGATGCTGCCTTTG  
AGCCTGTCTACTGGCTGGTAGACAACGTGATCCGCTGGTTTGGAGTGGTGTTCGTGGTCCTGG  
TGATCGTGCTGACAGGCTCCATTGTAGCTATCGCCTACCTGTGTGTCCTGCCTCTCATCCTCC  
GAACCTACTCAGTGCCACGACTCTGCTGGCATTCTTCTATAGCCACTGGAATCTGATCCTGA  
TTGTCTTCCACTACTACCAGGCCATCACCCTCCGCCTGGGTACCCACCCAGGGCAGGAATG  
ATATCGCCACCGTCTCCATCTGTAAGAAGTGCATTTACCCCAAGCCAGCCGAACACACCACT  
GCAGCATCTGCAACAGGTGTGTGCTGAAGATGGATCACCCTGCCCTGGCTAAACAATTGTG  
TGGGCCACTATAACCATCGGTACTTCTTCTCTTTCTGCTTTTTTCATGACTCTGGGCTGTGTCT  
ACTGCAGCTATGGAAGTTGGGACCTTTTCCGGGAGGCTTATGCTGCCATTGAGACTTATCACC  
AGACCCACCCACCCACCTTCTCCTTTCGAGAAAGGATGACTCACAAGAGTCTTGTCTACCTCT  
GGTTCCTGTGCAGTTCTGTGGCACTTGCCCTGGGTGCCCTAACTGTATGGCATGCTGTTCTCA  
TCAGTCGAGGTGAGACTAGCATCGAAAGGCACATCAACAAGAAGGAGAGACGTGCGCTACAGG  
CCAAGGGCAGAGTATTTAGGAATCCTTACAACTACGGCTGCTTGGACAACCTGGAAGGTATTCC  
TGGGTGTGGATACAGGAAGGCACTGGCTTACTCGGGTGCTCTTACCTTCTAGTCACTTGCCCC  
ATGGGAATGGAATGAGCTGGGAGCCCCCTCCCTGGGTGACTGCTCACTCAGCCTCTGTGATGG  
CAGT**GTGA**GCTGGACTGTGTCAGCCACGACTCGAGCACTCATTCTGCTCCCTATGTTATTTCA  
AGGGCCTCCAAGGGCAGCTTTTCTCAGAATCCTTGATCAAAAAGAGCCAGTGGGCCTGCCTTA  
GGGTACCATGCAGGACAATTCAAGGACCAGCCTTTTTTACCACTGCAGAAGAAAGACACAATGT  
GGAGAAATCTTAGGACTGACATCCCTTTACTCAGGCAAACAGAAGTTCCAACCCAGACTAGG  
GGTCAGGCAGCTAGCTACCTACCTTGCCAGTGCTGACCCGGACCTCCTCCAGGATACAGCAC  
TGGAGTTGGCCACCACCTCTTCTACTTGCTGTCTGAAAAACACCTGACTAGTACAGCTGAGA  
TCTTGGCTTCTCAACAGGGCAAAGATACCAGGCCTGCTGCTGAGGTCACTGCCACTTCTCACA  
TGCTGCTTAAGGGAGCACAAATAAAGGTATTGCATTTTTTAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAA

**FIGURE 252**

MRGQRSLLLGPRLCLRLLLLLGYRRRCPLLRLGLVQRWRYGKVCLRSLLYNSFGGSDTAVDA  
AFEPVYWLVDNVIRWFGVVFVVLVIVLTGSIVAIAIAYLCVLPLILRTYSVPRLCWHFFYSHWNL  
ILIVFHYYQAITTPPGYPPQGRNDIATVSICKKCIYPKPARTHHCSICNRCVLKMDHHCPWLN  
NCVGHYNHRYFFSFCFFMTLGCVYCSYGSWDLFREAYAAIETYHQTPPPTFSFRERMTHKSLV  
YLWFLCSSVALALGALTVWHAVLISRGETSIERHINKKERRRLQAKGRVFRNPYNYGCLDNWK  
VFLGVDTGRHWLTRVLLPSSHLPHGNGMSWEPPPWVTAHSASVMAV

**Important features:****Transmembrane domain:**

amino acids 88-100,202-216,254-274

**N-myristoylation sites:**

amino acids 55-61,56-62,92-98,210-216,309-315,319-325,340-346

**Prokaryotic membrane lipoprotein lipid attachment site:**

amino acids 201-212



GATCAAGGCGCCTTCCTTTCCCTTCTCCTCTCCCTACTTGGCCTTTGCCCTAAGCCAAGACCTGGCCATCAGCCTGCGC  
TGCAGGGGCGCTGCAGAGCCAGCTGCACCTTTTTTCAGGTATGGGGGAGGGCCAGGCACCATGAAAGCCAGTGTGGGTG  
GCCACCTTCTGTGGATGCTACTGCTGGTGCCAGGCTGGGGGCCGCCGGAAGGGGTCCCCAGAAGAGGCCTCC  
TTCTACTATGGAACCTTCCCTCTTGGCTTCTCCTGGGGCGTGGGCAGTTCTGCCTACCAGACGGAGGGCGCCTGG  
GACCAGGACGGGAAAGGGCCTAGCATCTGGGACGTCTTCACACACAGTGGGAAGGGGAAAGTGCTTGGGAATGAG  
ACGGCAGATGTAGCCTGTGACGGCTACTACAAGGTCCAGGAGGACATCATTTCTGCTGAGGGAACTGCACGTCAAC  
CACTACCGATTCTCCCTGTCTTGGCCCCGGCTCCTGCCACAGGCATCCGAGCCGAGCAGGTGAACAAGAAGGGA  
ATCGAATTCTACAGTGATCTTATCGATGCCCTTCTGAGCAGCAACATCACTCCCATCGTGACCTTGCACCACTGG  
GATCTGCCACAGCTGCTCCAGGTCAAATACGGTGGGTGGCAGAATGTGAGCATGGCCAACTACTTCAGAGACTAC  
GCCAACCTGTGCTTTGAGGCCTTTGGGGACCGTGTGAAGCACTGGATCACGTTCACTGATCCTCGGGCAATGGCA  
GAAAAAGGCTATGAGACGGGCCACCATGCGCCGGGCCTGAAGCTCCGCGGCACCGGCCTGTACAAGGCAGCACAC  
CACATCATTAAGGCCCACGCCAAAACCTGGCATTCTTATAACACCACGTGGCGCAGCAAGCAGCAAGGTCTGGTG  
GGAATTTCACTGAACTGTGACTGGGGGGAACCTGTGGACATTAGTAACCCCAAGGACCTAGAGGCTGCCGAGAGA  
TACCTACAGTTCTGTCTGGGCTGGTTTGCCAACCCCATTTATGCCGTGACTACCCCCAAGTCATGAAGGACTAC  
ATTGGAAGAAAGAGTGCAGAGCAAGGCCTGGAGATGTCGAGGTTACCGGTGTTCTCACTCCAGGAGAAGAGCTAC  
ATTAAAGGCACATCCGATTTCTTGGGATTAGGTCATTTTACTACTCGGTACATCACGGAAAGGAACTACCCCTCC  
CGCCAGGGGCCAGCTACCAGAACGATCGTGACTTGATAGAGCTGGTTGACCCAACTGGCCAGATCTGGGGTCT  
AAATGGCTATATTCTGTGCCATGGGGATTTAGGAGGCTCCTTAACCTTTGCTCAGACTCAATACGGTGATCCTCCC  
ATATATGTGATGGAAAATGGAGCATCTCAAAAATTCCACTGTACTCAATTATGTGATGAGTGGAGAATTCAATAC  
CTTAAAGGATACATAAATGAAATGCTAAAAGCTATAAAAGATGGTGCTAATATAAAGGGGTATACTTCCTGGTCT  
CTGTTGGATAAGTTTGAATGGGAGAAAGGATACTCAGATAGATATGGATTCTACTATGTTGAATTTAACGACAGA  
AATAAGCCTCGCTATCCAAAGGCTTCAGTTCAATATTACAAGAAGATTATCATTGCCAATGGGTTTCCCAATCCA  
AGAGAGGTGGAAAGTTGGTACCTCAAAGCTTTGGAACTTGCTCTATCAACAATCAGATGCTTGCTGCAGAGCCT  
TTGCTAAGTCACATGCAAATGGTTACGGAGATCGTGGTACCCACTGTCTGCTCCCTCTGTGTCCTCATCACTGCT  
GTTCTACTAATGCTCCTCCTGAGGAGGCAGAGCTGAGACAGGATTATCAATTTTGGAGCTTCATAAGAGAATCTT  
CAGGATCTTCCTCCCTTTTCTGCTTTGAGGGTTTCCATACATTGCTGTTTTTCAGGTTCTACAATAATTACCTTTT  
TTTCTCTTTCTCTTTTTTGGCTTGTGCTGGGATTTAAGAATTAGAAAAATAAAAAATAAGCAGAAATTA

**FIGURE 254**

MKPVVWATLLWMLLLVPRLLGAARKGSPEEASFYYGTFPLGFSWGVGSSAYQTEGAWDQDGKGPSIWDVFTHSGKG  
 KVLGNETADVACDGYKQVEDIILLRELHVNHYRFSLSWPRLLPTGIRAEQVNKKGIEFYSDLIDALLSSNITPI  
 VTLHHWDLPQLLQVKYGGWQNVSMANYFRDYANLCFEAFGDRVKHWITFSDPRAMAEKGYETGHHAPGLKLRGTG  
 LYKAAHHIIKAHAKTWHSYNTTWRSKQQGLVGISLNCDWGEPVDISNPKDLEAAERYLQFCLGWGFANPIYAGDYP  
 QVMKDYIGRKSAEQGLEMSRLPVFSLQEKSYIKGTSDFLGLGHFTTRYITERNYPSRQGPSYQNDRLIELVDPN  
 WPDLGSKWLYSVPWGFERRLLNFAQTQYGDPIIYVMENGASQKFHCTQLCDEWRIQYLKGYINEMLKAIKDGANIK  
 GYTSWSLLDKFEWEKGYSDRYGFYYVEFNDRNKPRYPKASVQYYKKIIANGFPNPREVESWYLKALETCSINNQ  
 MLAAEPLLSHMQMVTETIVVPTVCSLCVLITAVLLMLLLRRQS

**Important features:****Signal peptide:**

amino acids 1-21

**Transmembrane domain:**

amino acids 541-558

**N-glycosylation sites:**

amino acids 80-84,171-175,245-249

**Glycosaminoglycan attachment site:**

amino acids 72-76

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

amino acids 23-27,564-568

**Tyrosine kinase phosphorylation sites:**

amino acids 203-211,347-355,460-468,507-514

**N-myristoylation sites:**

amino acids 44-50,79-85,167-173,225-231,257-263,315-321

**Amidation site:**

amino acids 307-311

**Glycosyl hydrolases family 1 active site:**

amino acids 407-416

**Glycosyl hydrolases family 1 N-terminal signature:**

amino acids 41-56

**Motif name Glycosyl hydrolases family:**

amino acids 37- 67

**FIGURE 255**

CGCGAAG**ATG**CGAAAGGTGGTTTTTGATCACCGGGGCTAGCAGTGGCATTGGCCTGGCCCTCTG  
CAAGCGGCTGCTGGCGGAAGATGATGAGCTTCATCTGTGTTTGGCGTGCAGGAACATGAGCAA  
GGCAGAAGCTGTCTGTGCTGCTCTGCTGGCCTCTCACCCCACTGCTGAGGTCACCATTGTCCA  
GGTGGATGTCAGCAACCTGCAGTCGGTCTTCCGGGCCTCCAAGGAACCTAAGCAAAGGTTTCA  
GAGATTAGACTGTATATATCTAAATGCTGGGATCATGCCTAATCCACAACCTAAATATCAAAGC  
ACTTTTCTTTGGCCTCTTTTCAAGAAAAGTGATTCATATGTTCTCCACAGCTGAAGGCCTGCT  
GACCCAGGGTGATAAGATCACTGCTGATGGACTTCAGGAGGTGTTTGAGACCAATGTCTTTGG  
CCATTTTATCCTGATTCGGGAACCTGGAGCCTCTCCTCTGTACAGTGACAATCCATCTCAGCT  
CATCTGGACATCATCTCGCAGTGCAAGGAAATCTAATTTAGCCTCGAGGACTTCAGCACAG  
CAAAGGCAAGGAACCCTACAGCTCTTCCAAATATGCCACTGACCTTTTGAGTGTGGCTTTGAA  
CAGGAACCTCAACCAGCAGGGTCTCTATTCCAATGTGGCCTGTCCAGGTACAGCATTGACCAA  
TTTGACATATGGAATTCTGCCTCCGTTTATATGGACGCTGTTGATGCCGGCAATATTGCTACT  
TCGCTTTTTTTGCAAATGCATTCACCTTTGACACCATATAATGGAACAGAAGCTCTGGTATGGCT  
TTTCCACCAAAAAGCCTGAATCTCTCAATCCTCTGATCAAATATCTGAGTGCCACCACTGGCTT  
TGGAAGAAATTATATTATGACCCAGAAGATGGACCTAGATGAAGACACTGCTGAAAAATTTTA  
TCAAAAGTTACTGGAACCTGGAAGACACATTAGGGTCACTATTCAAAAAACAGATAATCAGGC  
CAGGCTCAGTGGCTCATGCCTA**TAA**TTCCAGCACTTTGGGAGGCCAAGGCAGAAGGATCACTT  
GAGACCAGGAGTTCAAGACCAGCCTGAGAAACATAGTGAGCCCTTGTCTCTACAAAAAGAAAT  
AAAAATAATAGCTGGGTGTGGTGGCATGCGCATGTAGTCCCAGCTACTCAGAAGGATGAGGTG  
GGAGGATCTCTTGAGGCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCACTGCACTCC  
AGCCTGGGTGACAGCGAGACCCTGTCTCAAAATATGTATATATTTAATATATATATAAAACCA  
GAGCTGACAATGACACTCTGGAACATTGCATACCTTCTGTACATTCTGGGGTACATGGATTTT  
TACTGAGTTGGATAATATGCATTTGTAATAAACTATGAACTATGAA

## **FIGURE 256**

MRKVVLITGASSGIGLALCKRLLAEDDELHLCLACRNMSKAEAVCAALLASHPTAEVTIVQVD  
VSNLQSVFRASKELKQRFQRLDCIYLNAGIMPNPQLNIKALFFGLFSRKVIHMFSTAEGLLTQ  
GDKITADGLQEVFETNVFGHFILIRELEPLLCHSDNPSQLIWTSSRSARKSNFSLEDFQHSGK  
KEPYSSSKYATDLLSVALNRNFNQQGLYSNVACPGTALTNLTYGILPPFIWTLMPAILLLRF  
FANAFTLTPYNGTEALVWLFHQKPESLNPLIKYLSATTGFGRNYIMTQKMDLDEDTAEKFYQK  
LLELEKHIRVTIQKTDNQARLSGSCL

**Important features:**

**Transmembrane domain:**

amino acids 234-254

**N-glycosylation sites:**

amino acids 37-41,178-182,229-233,263-267

**Glycosaminoglycan attachment site:**

amino acids 12-16

**N-myristoylation sites:**

amino acids 9-15,13-19,15-21,215-221,224-230

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332001>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 257**

CGGACGCGTGGGGCCGT**ATG**CGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCC  
 CCAGCCCCCTGCTCCTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGA  
 AGACCCGCCAGGTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTC  
 ATATACGGGCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGG  
 CAGTGGTAATGGTGGCCACCAACACCCCCACAGCACCTGAGCATCAACTGGAGCCTCCTGC  
 TATCCCCTGAGCCCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCTG  
 CCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCAAAGC  
 CTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATCACTGATT  
 CATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGACCCTACCAGGA  
 CTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCCAGCCGACCAGCCC  
 AACCCCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTGGCCCTGATTGGAGCCT  
 CTCCCCGGGGAAACCGTTCCCTGTTTTGGGCTGGAGGTAGCCACATTGGGCCAGGGCCCTGACT  
 GCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATATGCACCGGCCGTCTTCCAGTTGG  
 ACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCACAGTGGCGACCAGTGGCTTACTCCC  
 AGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGCCAAGCTTCCCCTCTTCATCCTGCCTTAG  
 CATACTCTCTTCCCCAGTCACCCATTGTCCGAGCCTTCTTTGGGTCCCAGAATAACTTCTGTG  
 CCTTCAATCTGACGTTTCGGGGCTTCCACAGGCCCTGGCTATTGGGACCAACACTACCTCAGCT  
 GGTCGATGCTCCTGGGTGTGGGCTTCCCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCA  
 TCATGGCAGTGGCCCTGGGTGCCCCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGC  
 TGCACCACAAGAAGTACTCAGAGTACCAGTCCATAAAT**TAA**GGCCCGCTCTCTGGAGGGAAGG  
 ACATTACTGAACCTGTCTTGCTGTGCCTCGAACTCTGGAGGTTGGAGCATCAAGTTCAGCC  
 GGCCCCCTTCACTCCCCCATCTTGCTTTTCTGTGGAACCTCAGAGGCCAGCCTCGACTTCCTGG  
 AGACCCCCAGGTGGGGCTTCCCTCATACTTTGTTGGGGGACTTTGGAGGCGGGCAGGGGACAG  
 GGCTATTGATAAGGTCCCCTTGGTGTTCCTTGCATCTCCACACATTTCCCTTGGATGGG  
 ACTTGCAGGCCTAAATGAGAGGCATTCTGACTGGTTGGCTGCCCTGGAAGGCAAGAAAATAGA  
 TTTATTTTTTTTTCACAGGGGAAAAAAAAAAAA

## **FIGURE 258**

MRGSVECTWGWGHCAPSPLLLWTLFFFAPFGLLGEKTRQVSLEVI PNWLGPLQNLLHIRAVG  
TNSTLHYVWSSLGPLAVVMVATNTPHSTLSINWSLLLSPEPDGGLMVLPKDSIQFSSALVFTR  
LLEFDSTNVSDTAAKPLGRPYPPYSLADFSWNNITDSLDPATLSATFQGHMNDPTRTFANGS  
LAFRVQAQFSRSSRPAQPPRLLHTADTCQLEVALIGASPRGNRSLFGLEVATLGQGPDCPSMQE  
QHSIDDEYAPAVFQLDQLLWGS LPSGFAQWRPVAYSQKPGGRESALPCQASPLHPALAYSLPQ  
SPIVRAFFGSQNNFCAFNLTFGASTGPGYWDQHYLSWSMLLGVGFPPVDGLSPLVLGIMAVAL  
GAPGLMLLGGGLVLLLHHKKYSEYQSIN

### **Important features:**

#### **Signal peptide:**

amino acids 1-35

#### **Transmembrane domain:**

amino acids 365-386

#### **N-glycosylation sites:**

amino acids 65-69, 95-99, 134-138, 159-163, 187-191, 230-234, 333-337

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 397-401

#### **N-myristoylation sites:**

amino acids 3-9, 63-69, 235-241, 273-279, 292-298, 324-330

#### **Leucine zipper pattern:**

amino acids 371-393

**FIGURE 259**

**FIGURE 260**

MVSKALLRLVSAVNRRRMKLLLGIALLAYVASVWGNFVNMRSIQENGELKIESKIEEMVEPLR  
EKIRDLEKSFTQKYPPVKFLSEKDRKRILITGGAGFVGSHLTDKLMMDGHEVTVVDNFFTGRK  
RNVEHWIGHENFELINHADVVEPLYIEVDQIYHLASPPNYMYPNPIKTLKTNTIGTLNMLGL  
AKRVGARLLLASTSEVYGDPEVHPQSEDYWGHVNPIGPRACYDEGKRVAETMCYAYMKQEGVE  
VRVARIFNTFGPRMHMNDGRVVSFILQALQGEPLTVYGSQSQTRAFQYVSDLVNGLVNLMNS  
NVSSPVNLGNPEEHTILEFAQLIKNLVSGSGSEIQFLSEAQDDPQKRKPDIKKAKMLGWEPVV  
PLEEGLNKAIHYFRKELEYQANNQYIPKPKPARIKKGRTRHS

**Important features:****Signal peptide:**

amino acids 1-32

**N-glycosylation site:**

amino acids 316-320

**Tyrosine kinase phosphorylation site:**

amino acids 235-244

**N-myristoylation sites:**

amino acids 35-41,101-107,383-389

**Amidation sites:**

amino acids 123-127,233-237



**FIGURE 261**

GCGTGGTGCGGGGGCGTGGGGAAATCGGGTTGCCCCAGCCGTTACTGGTCCGCGCAGTCAGGG  
CATCCTCCGCATCCTCCACATCCTTCC**ATG**GCTCTGAAGAATAAATTCAGTTGTTTATGGATC  
TTGGGTCTGTGTTTGGTAGCCACTACATCTTCCAAAATCCCATCCATCACTGACCCACACTTT  
ATAGACAACCTGCATAGAAGCCCACAACGAATGGCGTGGCAAAGTCAACCCTCCCGCGGCCGAC  
ATGAAATACATGATTTGGGATAAAGGTTTAGCAAAGATGGCTAAAGCATGGGCAAACCAGTGC  
AAATTTGAACATAATGACTGTTTGGATAAATCATATAAATGCTATGCAGCTTTTGAATATGTT  
GGAGAAAATATCTGGTTAGGTGGAATAAAGTCATTCACACCAAGACATGCCATTACGGCTTGG  
TATAATGAAACCCAATTTTATGATTTTGATAGTCTATCATGCTCCAGAGTCTGTGGCCATTAT  
ACACAGTTAGTTTGGGCCAATTCATTTTATGTCGGTTGTGCAGTTGCAATGTGTCCTAACCTT  
GGGGGAGCTTCAACTGCAATATTTGTATGCAACTACGGACCTGCAGGAAATTTTGCAAATATG  
CCTCCTTACGCAAGAGGAGAATCTTGCTCTCTCTGCTCAAAAGAAGAGAAATGTGTAAAGAAC  
CTCTGCAGGACTCCACAACCTTATTATACCTAACCAAAATCCATTTCTGAAGCCAACGGGGAGA  
GCACCTCAGCAGACAGCCTTTAATCCATTCAGCTTAGGTTTTCTTCTTCTGAGAATCTTT**TAA**  
TGTCATTTATATACAAAAGAAATTCTCAAATGTTAAATAAAGGAATAGTTTATTGCTTAATA

## **FIGURE 262**

MALKNKFSCWLWILGLCLVATTSSKIP SITDPHFIDNCIEAHNEWRGKVNPPAADMKYMIWDKG  
LAKMAKAWANQCKFEHNDCLDKSYKCYAAFEYVGENIWLGGIKSFTPRHAITAWYNETQFYDF  
DSLSCSRVCGHYTQLVWANSFYVGCAMCPNLGGASTAIFVCNYGPAGNFANMPPYARGESC  
SLCSKEEKCCKVKNLCRTPQLIIPNQNPFLKPTGRAPQQTAFNPFSLGFLLLRIF

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **N-glycosylation site:**

amino acids 119-123

#### **N-myristoylation sites:**

amino acids 103-109,150-156,160-166,161-167,175-181

#### **Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1:**

amino acids 136-156

#### **Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2:**

amino acids 166-178

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 263

CGCCCTCCGACCCCGCCCGCGGCATTTGTGGGATCTGTCTGGCTTGTCTCAGGTGGTGGAGGAAA  
AGGCGCTCCGTCATGGGGATCCAGACGAGCCCCGTCCTGCTGGCCTCCCTGGGGGTGGGGCTG  
GTCACTCTGCTCGGCCTGGCTGTGGGCTCCTACTTGGTTTCGGAGGTCCCGCCGGCCTCAGGTC  
ACTCTCCTGGACCCCAATGAAAAGTACCTGCTACGACTGCTAGACAAGACGACTGTGAGCCAC  
AACACCAAGAGGTTCCGCTTTTGGCCTGCCCACCGCCCACCACACTCTGGGGCTGCCTGTGGGG  
AAACATATCTACCTCTCCACCCGAATTGATGGCAGCCTGGTCATCAGGCCATACACTCCTGTCT  
ACCAGTGATGAGGATCAAGGCTATGTGGATCTTGTTCATCAAGGTCTACCTGAAGGGTGTGCAC  
CCCAAATTTCTGAGGGAGGGAAGATGTCTCAGTACCTGGATAGCCTGAAGGTTGGGGATGTG  
GTGGAGTTTCGGGGGGCCAAGCGGGTTGCTCACTTACACTGGAAAAGGGCATTTTAACATTTCAG  
CCCAACAAGAAATCTCCACCAGAACCCCGAGTGGCGAAGAACTGGGAATGATTGCCGGCGGG  
ACAGGAATCACCCCAATGCTACAGCTGATCCGGGGCCATCCTGAAAGTCCCTGAAGATCCAACC  
CAGTGCTTTCTGCTTTTTTGCCAACCAGACAGAAAAGGATATCATCTTGCGGGAGGACTTAGAG  
GAACTGCAGGCCCCGCTATCCCAATCGCTTTAAGCTCTGGTTCACTCTGGATCATCCCCAAAA  
GATTGGGCCTACAGCAAGGGCTTTGTGACTGCCGACATGATCCGGGAACACCTGCCCGCTCCA  
GGGGATGATGTGCTGGTACTGCTTTGTGGGCCACCCCCAATGGTGCAGCTGGCCTGCCATCCC  
AACTTGGACAACTGGGCTACTCACAAAAGATGCGATTACCTACTTGAGCATCCTCCAGCTTC  
CCTGGTGCTGTTTCGCTGCAGTTGTTCCCCATCAGTACTCAAGCACTATAAGCCTTAGATTCTT  
TTCTCAGAGTTTCAGGTTTTTTTCAGTTACATCTAGAGCTGAAATCTGGATAGTACCTGCAGG  
AACAAATATTCTGTAGCCATGGAAGAGGGCAAGGCTCAGTCACTCCTTGGATGGCCTCCTAAA  
TCTCCCCGTGGCAACAGGTCCAGGAGAGGGCCATGGAGCAGTCTCTTCCATGGAGTAAGAAGG  
AAGGGAGCATGTACGCTTGGTCCAAGATTGGCTAGTTCCTTGATAGCATCTTACTCTCACCTT  
CTTTGTGTCTGTGATGAAAGGAACAGTCTGTGCAATGGGTTTTACTTAAACTTCACTGTTCAA  
CCTATGAGCAAATCTGTATGTGTGAGTATAAGTTGAGCATAGCATACTTCCAGAGGTGGTNTT  
ATGGAGATGGCAAGAAAGGAGGAAATGATTTCTTCAGATNTCAAAGGAGTCTGAAATATCATA  
TTTCTGTGTGTGTCTCTCTCAGCCCCTGCCCAGGCTAGAGGGAAACAGCTACTGATAATCGAA  
AACTGCTGTTTGTGGCANGAACCCCTGGCTGTGCAAATAAATGGGGCTGAGGCCCTGTGTGA  
TATTGAAGA

## **FIGURE 264**

MGIQTSPVLLASLGVLVTLLGLAVGSYLVRRSRRPQVTLLDPNEKYLLRLLDKTTVSHNTR  
FRFALPTAHHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSDEDQGYVDLVIKVYLKGVHPKFP  
EGGKMSQYLDLKVGDVVEFRGPSGLLTYTGKGHFNIQPNKKSPPEPRVAKKLGMIAGGTGIT  
PMLQLIRAILKVPEDPTQCFLLEFANQTEKDIILREDLEELQARYPNRFKLWFTLDHPPKDWAY  
SKGFVTADMIREHLPA PGDDVLVLLCGPPPMVQLACHPNLDKLGYSQKMRFTY

### **Important features:**

#### **Signal peptide:**

amino acids 1-26

#### **N-glycosylation site:**

amino acids 214-218

#### **N-myristoylation sites:**

amino acids 22-28, 76-82, 128-134, 180-186

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 265**

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAGA  
ACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTATAG  
AATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTCGGT  
TCTATCGATAATCTCAGCACCAAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCCGCCTCAGG  
CTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCAATGCC  
TCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCCAGGAAC  
AGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGACCATCTAC  
AGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGATGAGCAGA  
AGATACCTCTGCATGGATTTTCAGAGGCAACATTTTTTGGATCACACTATTTTCGACCCGGAGAAC  
TGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCCTCAGTATCAC  
TTCCTGGTTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACCCACCCCGTAC  
TCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACCCCATACCACGG  
CGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCTGAACGTGCTGAAGCCCCGG  
GCCCCGATGACCCCGGCCCCGGCCTCCTGTTTACAGGAGCTCCCGAGCGCCGAGGACAACAGC  
CCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTTCGAGTGAACACGCACGCTGGGGGA  
ACGGGCCCCGGAAGGCTGCCGCCCCTTCGCCAAGTTCATC**TAG**GGTTCGCTGG

**FIGURE 266**

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWSGGLIHLYTATARNSYHLQIHKNGHVDG  
APHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENG YDV  
YHSPQYHFLVSLGRAKRAFLPGMNPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSERDP  
LNVLKPRARMT PAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFAKFI

**Important features:****Signal peptide:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 175-179

**N-myristoylation site.**

amino acids 33-39, 100-106, 225-231, 229-235

**HBGF/FGF family proteins**

amino acids 73-124

**FIGURE 267**

GGCTGAGGGGAGGCCCCGAGCCTTTCTGGGGCCTGGGGGATCCTCTTGCACTGGTGGGTGGAGAGAAGCGCCTGC  
 AGCCAACCAGGGTCAGGCTGTGCTCACAGTTTCCTCTGGCGGCATGTAAAGGCTCCACAAAGGAGTTGGGAGTTC  
 AAATGAGGCTGCTGCGGACGGCCTGAGGATGGACCCCAAGCCCTGGACCTGCCGAGCGTGGCACTGAGGCAGCGG  
 CTGACGCTACTGTGAGGGAAAGAAGGTTGTGAGCAGCCCCGAGGACCCCTGGCCAGCCCTGGCCCCAGCCTCTG  
 CCGGAGCCCTCTGTGGAGGCAGAGCCAGTGGAGCCCAGTGAAGCAGGGCTGCTTGGCAGCCACCGGCCTGCAACT  
 CAGGAACCCCTCCAGAGGCCATGGACAGGCTGCCCCGCTGACGGCCAGGGTGAAGCATGTGAGGAGCCGCCCCGG  
 AGCCAAGCAGGAGGGAAGAGGCTTTTCATAGATTCTATTACAAAGAATAACCACCATTTTGCAAGGACCATGAGG  
 CCACTGTGCGTGACATGCTGGTGGCTCGGACTGCTGGCTGCCATGGGAGCTGTTGCAGGCCAGGAGGACGGTTTT  
 GAGGGCACTGAGGAGGGCTCGCCAAGAGAGTTCATTTACCTAAACAGGTACAAGCGGGCGGGCGAGTCCCAGGAC  
 AAGTGCACCTACACCTTCATTGTGCCCCAGCAGCGGGTCACGGGTGCCATCTGCGTCAACTCCAAGGAGCCTGAG  
 GTGCTTCTGGAGAACCGAGTGCATAAGCAGGAGCTAGAGCTGCTCAACAATGAGCTGCTCAAGCAGAAGCGGCAG  
 ATCGAGACGCTGCAGCAGCTGGTGGAGGTGGACGGCGGCATTGTGAGCGAGGTGAAGCTGCTGCGCAAGGAGAGC  
 CGCAACATGAACTCGCGGGTCACGCAGCTCTACATGCAGCTCCTGCACGAGATCATCCGCAAGCGGGACAACGCG  
 TTGGAGCTCTCCCAGCTGGAGAACAGGATCCTGAACCAGACAGCCGACATGCTGCAGCTGGCCAGCAAGTACAAG  
 GACCTGGAGCACAAGTACCAGCACCTGGCCACACTGGCCACAACCAATCAGAGATCATCGCGCAGCTTGAGGAG  
 CACTGCCAGAGGGTGCCCTCGGCCAGGCCCCGTCCCCAGCCACCCCCCGCTGCCCCGCCCCGGGTCTACCAACCA  
 CCCACCTACAACCGCATCATCAACCAGATCTCTACCAACGAGATCCAGAGTGACCAGAACCTGAAGGTGCTGCCA  
 CCCCCCTCTGCCCCTATGCCCCTCTCACCAGCCTCCCCTCTTCCACCGACAAGCCGTGCGGGCCCATGGAGAGAC  
 TGCTTGACAGGCCCTGGAGGATGGCCACGACACCAGCTCCATCTACCTGGTGAAGCCGGAGAACCAACCGCCTC  
 ATGCAGGTGTGGTGCGACCAGAGACACGACCCCGGGGGCTGGACCGTCATCCAGAGACGCCTGGATGGCTCTGTT  
 AACTTCTTCAGGAACCTGGGAGACGTACAAGCAAGGGTTTGGGAACATTGACGGCGAATACTGGCTGGGCTGGAG  
 AACATTTACTGGCTGACGAACCAAGGCAACTACAACTCCTGGTGACCATGGAGGACTGGTCCGGCCGCAAAGTC  
 TTTGCAGAATACGCCAGTTTCCGCCTGGAACCTGAGAGCGAGTATTATAAGCTGCGGCTGGGGCGCTACCATGGC  
 AATGCGGGTGACTCCTTTACATGGCACAACGGCAAGCAGTTCACCACCCTGGACAGAGATCATGATGTCTACACA  
 GGAACTGTGCCCCTACCAGAAGGGAGGCTGGTGGTATAACGCCTGTGCCCCTCCAACCTCAACGGGGTCTGG  
 TACCGCGGGGGCCATTACCGGAGCCGCTACCAGGACGGAGTCTACTGGGCTGAGTTCGAGGAGGCTCTTACTCA  
 CTCAAGAAAGTGGTGATGATGATCCGACCGAACCCCAACACCTTCCACTTAAGCCAGCTCCCCCTCCTGACCTCTC  
 GTGGCCATTGCCAGGAGCCCACCCTGGTCACGCTGGCCACAGCACAAAGAACAACCTCCTCACCAGTTCATCCTGA  
 GGCTGGGAGGACCGGGATGCTGGATTCTGTTTTCCGAAGTCACTGCAGCGGATGATGGAACCTGAATCGATACGGT  
 GTTTTCTGTCCCTCCTACTTTTCCTTCACACCAGACAGCCCCCTCATGTCTCCAGGACAGGACAGGACTACAGACAA  
 CTCTTTCTTTAAATAAATTAAGTCTCTACAATAAAAAAA

**FIGURE 268**

MRPLCVTCWWLGLLAAMGAVAGQEDGFEGTEEGSPREFIYLNRYKRAGESQDKCTYTFIVPQQ  
 RVTGAICVNSKEPEVLLLENRVHKQELELLNNELLKQKRQIETLQQQLEVDGGIVSEVKLLRKE  
 SRNMNSRVTQLYMQLLHEIIRKRDNALELSQLENRILNQTADMLQLASKYKDLEHKYQHLATL  
 AHNQSEIIAQLEEHQCRVPSARPVPQPPPAAPPRVYQPPTYNRIINQISTNEIQSDQNLKVLP  
 PPLPTMPTLTSLPSSTDKPSGPWRDCLQALEDGHDTSSIYLVKPENTNRLMQVWCDQRHDPGG  
 WTVIQRRLDGSVNFFRNWETKQGFQGNIDGEYWLGLENIYWLTNQGNKLLVTMEDWSGRKVF  
 AEYASFRLEPESEYYKLRLGRYHGNAGDSFTWHNGKQFTTLDRDHDVYTGNCAYQKGGWWYN  
 ACAHSNLNGVWYRGGHYRSRYQDGVYWAEFRGGSYSLKKVVMIRPNPNTFH

**Important features:****Signal peptide:**

amino acids 1-22

**N-glycosylation sites:**

amino acids 164-168, 192-196

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 124-128

**Tyrosine kinase phosphorylation sites:**

amino acids 177-184, 385-393, 385-394, 461-468

**N-myristoylation sites:**amino acids 12-18, 18-24, 22-28, 29-35, 114-120, 341-347, 465-471,  
473-479**Amidation site:**

amino acids 373-377

**Fibrinogen beta and gamma chains C-terminal domain signature:**

amino acids 438-451

**Fibrinogen beta and gamma chains C-terminal domain proteins:**

amino acids 305-343, 365-402, 411-424, 428-458

**Trehalase proteins:**

amino acids 275-292



CCCGAGCTGAGCGGATCCTCAC**ATG**ACTGTGATCCGATTCTTTCCAGCGGGCTTCTGCAACCAAGCGGGTCTTACCCCGGTCTCCGCGTCTCCAGTCTCGCACCTGGAACCCCAACGTCCCCGAGAGTCCCCGAATCCCCGCTCCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGACGGCCGGGGCAGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTCAGGGCGGACCCGTGCAGTCAAGTCGCCGCGCTTTGCGTCTTGGGACGAGATGAATGTCTGGCGCACGGACTCCTGCAGCTCGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGCAGTCAGCTGAGCGCGCTGGAGCGGGCCCTGAGCGCGTGCGGGTCCGCCTGTCAGGGAACCGAGGGGTCCACCGACCTCCCGTTAGCCCCTGAGAGCCGGGTGGACCCTGAGGTCTTACAGCCTGCAGACACAACCTCAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCAGCAGCAGCGGCACCTGGAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGACCACAAGCACCTAGACCATGAGGTGGCCAAGCCTGCCCGAAGAAAGAGGCTGCCCGAGATGGCCAGCCAGTTGACCCGGCTCACAA TGTCAGCCGCTGCACCGGTGCCCAGGGATTGCCAGGAGCTGTTCCAGGTTGGGGAGAGGCCAGAGTGGACTATTTGAAATCCAGCCTCAGGGGTCTCCGCCATTTTTTGGTGAACCTGCAAGATGACCTCAGATGGAGGCTGGACAGTAATTCAGAGGCGCCACGATGGCTCAGTGGACTTCAACCGGCCCTGGGAAGCCTACAAGCGGGGTTTTGGGGATCCCCACGGCGAGTTCTGGCTGGGTCTGGAGAA GGTGCATAGCATCACGGGGGACCGCAACAGCCGCTGGCCGTGCAGCTGCGGGACTGGGATGGCAACGCCGAGTTGCTGCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCA GCTCACTGCACCCGTGGCCGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGTACCCTTCTCCACTTGGGACCAGGATCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAGCCTCTCTGGAGGCTGGTGGTTTTGGCACCTGCAGCCATTCCAACCTCAACGGCCAGTACTTCCGCTCATCCCACAGCAGCGGCAGAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGGCCGCTACTACCCGCTGCAGGCCACCACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCC**TAG**CGTCCTGGCTGGGCCTGGTCCCAGGCCCACGAAAGACGGTGACTCTTGGCTCTGCCCGAGGATGTGGCCGTTCCCTGCCTGGGCAGGGGCTCCAAGGAGGGGCCATCTGGAAACTTGTGGACAGAGAA GAAGACCACGACTGGAGAAGCCCCCTTTCTGAGTGCAGGGGGGCTGCATGCGTTGCCTCCTGA GATCGAGGCTGCAGGATATGCTCAGACTCTAGAGGCGTGGACCAAGGGGCATGGAGCTTCACTCCTTGCTGGCCAGGGAGTTGGGGACTCAGAGGGACCCTTGGGGCCAGCCAGACTGGCCTCAA TGGCGGACTCAGTCACATTGACTGACGGGGACCAGGGCTTGTGTGGGTGAGAGCGCCCTCATGGTGCTGGTGCTGTTGTGTGTAGGTCCCCTGGGGACACAAGCAGGCGCCAATGGTATCTGGGC GGAGCTCACAGAGTTCTTGAATAAAAGCAACCTCAGAACAC

**FIGURE 270**

MTVIRFFPAASATKRVLPPVLRVSSPRTWNPVPEPRIPAPRLPKRMSGAPTAGAALMLCAA  
TAVLLSAQGGPVQSKSPREFASWDEMNVLAHGLLQLGQGLREHAERTRSQLSALERRLSACGSA  
CQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLFHKVAQQQRHLEKQHLRIQHLQS  
QFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSRLHRLPRDCQELFQVGERQSGLFELIQ  
PQGSPPFLVNCKMTSDGGWTVIQRRHDGSDVFNRPWEAYKAGFGDPHGEFWLGLEKVHSITGD  
RNSRLAVQLRDWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPVPSGLSVPFSTWDQ  
DHDLRRDKNCAKSLSGGWWFGTCSHSNLNGQYFRSIPQQRQKLKKGIFWKTRGRYYPLQATT  
MLIQPMAAEAAS

**Important features:****Signal peptide:**

Amino acids 1-13

**Transmembrane domain:**

Amino acids 53-70

**N-glycosylation site:**

Amino acids 224-228

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 46-50;118-122

**N-myristoylation sites:**

Amino acids 50-56;129-135;341-347;357-363

**Fibrinogen beta and gamma chains C-terminal domain signature:**

Amino acids 396-409

**FIGURE 271**

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGGG  
AACAAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGCTG  
CTGCTGCTGACCATGGCCTTGGCCGGAGGTTCTGGGGACCGCTTCGGCTGAAGCATTGACTCG  
GTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACACCTAC  
CCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTT  
GTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA  
TATCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTTCGCT  
GAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAACT  
CTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACCTCTTCATGG  
ACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCCAGAAATCCAG  
TACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAAGCAAAATGTCC  
TATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTCTTGAAGATGGAGAAAGTGATGGC  
TTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACCTCTTGTCCTCTCGGTG  
ATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGCAGTATGTTCCCTCT  
GAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAGCTAAACAGATATCCA  
GCTTCTTCTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGAAGCAGGGCCTCTACCT  
ACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTTAAAGACAAGTGTAATA  
GACATCTAAAATTCACCTCCTCATAGAGCTTTTAAATGGTTTCATTGGATATAGGCCTTAAG  
AAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

**FIGURE 272**

MAAPKGS LWVRTQLGLPPLLLLT MALAGGSGTASAEAFDSVLGDTASCHRA CQLTYPLHTYPK  
EEELYACQ RGCRLFSICQFVDDGIDLNR TKLECESACTEAYSQSDEQYACHLGCQNQLPFAEL  
RQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQS FITSSWTFYLQADDGKIVIFQSKPEIQYA  
PHLEQEPTNLRESSLSKMSY LQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVLSVMV  
LLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLN RYPASSLVVVRSKTEDHEEAGPLPTK  
VNLAHSEI

**Important features:****Signal peptide:**

amino acids 1-31

**Transmembrane domain:**

amino acids 241-260

**N-glycosylation site:**

amino acids 90-94

**N-myristoylation sites:**

amino acids 28-34,29-35,31-37,86-92

**FIGURE 273**

CCCACGCGTCCGAACCTCTCCAGCG**ATG**GGAGCCGCCCCGCTGCTGCCCCAACCTCACTCTGTG  
CTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACCGA  
CCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCACGT  
GCAGGTCACCGGGCGTTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGCTCAT  
AGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAGTACAT  
CTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTGCGTGTT  
CACGGAGATCGTGCTGGAGAACAATAACGGCCTTCCAGAACGCCCGGCACGAGGGCTGGTT  
CATGGCCTTCACGCGGCAGGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAACCAGCGCGA  
GGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCCAGCCGAGAAGCA  
GAAGCAGTTTCGAGTTTGTGGGCTCCGCCCCACCCGCCGACCAAGCGCACACGGCGGGCCCCA  
GCCCCTCACG**TAG**TCTGGGAGGCAGGGGGCAGCAGCCCCCTGGGCCGCCTCCCCACCCCTTTCC  
CTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCCGAGGGAGGACCC  
TGAGGGCCCGGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTGCCCCAGGGGCGGC  
TGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACTGAGTGTCACCCTGA  
TCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCCCGCTGAAAGGTCAGC  
GACTGAAGGCCTTGACAGACAACCGTCTGGAGGTGGCTGTCCTCAAAATCTGCTTCTCGGATCT  
CCCTCAGTCTGCCCCCAGCCCCCAAACCTCCTCCTGGCTAGACTGTAGGAAGGGACTTTTGTTT  
GTTTGTTTGTTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAATAGAGGGTTGTCCACTCCTCA  
CATTCCACGACCCAGGCCTGCACCCACCCCCAACTCCCAGCCCCGGAATAAAACCATTTTCC  
TGC

## **FIGURE 274**

MGAARLLPNLTLCLQLLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRIS  
ATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENN  
YTAFQONARHEGWFMATFTRQGRPRQASRSRQNRQEAHFQKRLYQGQLPFPNHAEKQKQFEFVGS  
APTRRTKRTRRPQPLT

### **Important features:**

#### **Signal peptide:**

Amino acids 1-22

#### **N-glycosylation site.**

amino acids 9-13, 126-130

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 60-64

#### **Casein kinase II phosphorylation site.**

amino acids 65-69

#### **Tyrosine kinase phosphorylation site.**

amino acids 39-48, 89-97

#### **N-myristoylation site.**

amino acids 69-75, 188-194

#### **Amidation site.**

amino acids 58-62

#### **HBGF/FGF family signature.**

amino acids 103-128

**FIGURE 275**

TATTTACCATATCAGATTACATTTCAGTCCTCAGCAAAATGAAGGGCTCCATTTTCACTCTGT  
TTTTATTCTCTGTCCTATTTGCCATCTCAGAAGTGCGGAGCAAGGAGTCTGTGAGACTCTGTG  
GGCTAGAATACATACGGACAGTCATCTATATCTGTGCTAGCTCCAGGTGGAGAAGGCATCTGG  
AGGGGATCCCTCAAGCTCAGCAAGCTGAGACAGGAACTCCTTCCAGCTCCCACATAAACGTG  
AGTTTTCTGAGGAAAATCCAGCGCAAAACCTTCCGAAGGTGGATGCCTCAGGGGAAGACCGTC  
TTTGGGGTGGACAGATGCCCCTGAAGAGCTTTGGAAGTCAAAGAAGCATTTCAGTGATGTCAA  
GACAAGATTTACAACTTTGTGTTGCACTGATGGCTGTTCCATGACTGATTTGAGTGCTCTTT  
GC**TAA**GACAAGAGCAAATACCCAATGGGTGGCAGAGCTTTATCACATGTTTAATTACAGTGTT  
TACTGCCTGGTAGAACACTAATATTGTGTTATTAAAATGATGGCTTTTGGGTAGGCCAAA  
TCTTTTCTAAAAGGTATAGCTGAGCGGTTGAAACCACAGTGATCTCTATTTTCTCCCTTTGCC  
AAGGTTAATGAACTGTTCTTTTCAAATTCTACTAATGCTTTGAAATTTCAAATGCTGCGCAA  
ATTGCAATAAAAATGCTATAAA

## **FIGURE 276**

MKGSIFTLFLFSVLFAISEVRSKESVRLCGLEYIRTVIYICASSRWRRLHLEGIPQAQQAETGN  
SFQLPHKREFSEENPAQNLPKVDASGEDRLWGGQMPTEELWKSCKHVSMSRQDLQTLCTDGC  
SMTDLSALC

### **Important features:**

#### **Signal sequence:**

amino acids 1-18

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site:**

amino acids 107-111

#### **N-myristoylation sites:**

amino acids 3-9, 52-58, 96-102, 125-131

#### **Insulin family signature:**

amino acids 121-136

#### **Insulin family proteins:**

amino acids 28-46

276/550





**FIGURE 278**

MKTFTWTLGVLFLLVDTGHCRCGGQFKIKKINQRRYPRATDGKEEAKKCAYTFLVPEQRITGP  
ICVNTKGQDASTIKDMITRMDLENLKDVLRSQKREIDVLQLVVDVDGNIVNEVKLLRKESRNM  
NSRVTQLYMQLLHEIIRKRDNSLELSQLENKILNVTTEMLKMATRYRELEVKYASLTDLVNNQ  
SVMITLLEEQCLRIFSRQDTHVSPPLVQVVPQHIPNSQQYTPGLLGGNEIQRDPGYPRDLMPP  
PDLATSPTKSPFKIPPVTFINEGPFKDCQQAKEAGHSVSGIYMIKPENSNGPMQLWCENSLDP  
GGWTVIQKRTDGSVNFFRNWENYKKGFGNIDGEYWLGLENIYMLSNQDNYKLLIELEDWSDKK  
VYAEYSSFRLEPESEFYRLRLGTYYQGNAGDSMMWHNGKQFTTLDRDKDMYAGNCAHFHKGWW  
YNACAHSNLNGVWYRGGHYRSKHQDGI FWA EYRGGSYSLRAVQMMIKPID

**Important features:****Signal sequence:**

Amino acids 1-23

**N-glycosylation sites:**

Amino acids 160-164;188-192

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 120-124

**Tyrosine kinase phosphorylation sites:**

Amino acids 173-180;387-396

**N-myristoylation sites:**

Amino acids 70-76;110-116;232-238,343-349;400-406;467-473;  
475-487

**Fibrinogen beta and gamma chains C-terminal domain signature:**

Amino acids 440-453

**FIGURE 279**

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGGC  
 CCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCCGCGCGGGAGCCGGACCGCCG  
 CCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAAGCC  
 CGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAGCGGA  
 GGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCAGCAAAGAAGCGGTGGTGGTGGGCGTCG  
 TGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCGAGCGCG  
 AGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGCGACAAAA  
 ACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAGAAGAAGAC  
 CAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACCACTTGCAGC  
 TGCAGGCGGATGGAACCATTTGATGGCACCAAGATGAGGACAGCACTTACACTCTGTTTAACC  
 TCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCTGTACTTGGCAA  
 TGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTACACCTGAGTGCAAATTCAAAGAAT  
 CAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAGCAGCAGTCAGGCC  
 GAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAACCATGTGAAGAAGA  
 ACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGTACAAGGAGCCATCAC  
 TGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACCAAGAGCAGAAGTGTCT  
 CTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAACG**TAG**CCAGTGAGGGCAA  
 AAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAATTCTTCTAGCAGTCCTTCA  
 CCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGGCAGAGTTCACTATTCTATC  
 TGCCATTAGACCTTCTTATCATCCATACTAAAGC

## **FIGURE 280**

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFSGSKRRRRRPE  
PQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLI PVGLRVVAIQGVQTKLYLAMN  
SEGYLYTSELF TPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVKKNK  
PAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNNGGKSMSHNEST

### **Important Features:**

#### **N-glycosylation site:**

Amino acids 242-246

#### **Glycosaminoglycan attachment sites:**

Amino acids 165-169, 218-222

#### **Tyrosine kinase phosphorylation site:**

Amino acids 93-100

#### **N-myristoylation sites:**

Amino acids 87-93, 231-237

#### **ATP/GTP-binding site motif A (P-loop):**

Amino acids 231-239

#### **HBGF/FGF family proteins:**

Amino acids 78-94, 102-153



## **FIGURE 282**

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRRPARPDRG  
PEPQLKGIVTKLFCRQGFYQLQANPDGSIQGTPEDTSSSFTHFNLI PVGLRVVTIQSAKLGHYMA  
MNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNRVKK  
TKAAAHFLPKLLEVAMYQEPSLHSVPEASPSPPAP

### **Important features:**

#### **Tyrosine kinase phosphorylation site:**

Amino acids 199-207

#### **N-myristoylation sites:**

Amino acids 54-60; 89-95; 131-137

#### **HBGF/FGF family signature:**

Amino acids 131-155



## **FIGURE 284**

MAAAIASGLIRQKRQAREQHWDPRPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLRR  
QDPQLKGIVTRLYCRQGYYLQMHPDGALDGTKDDSTNSTLFNLI PVGLRVVAIQGVKTGLYIA  
MNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNRVKK  
TKPAAHFLPKPLEVAMYREPSLHDVGETVPPKPGVTPSKSTSASAIMNGGKPVNKSSTT

### **Important features:**

#### **N-glycosylation sites:**

Amino acids 100-104, 242-246

#### **cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 28-32, 29-33

#### **Tyrosine kinase phosphorylation site:**

Amino acids 199-207

#### **N-myristoylation sites:**

Amino acids 38-44, 89-95, 118-124, 122-128, 222-228

#### **HBGF/FGF family proteins:**

Amino acids 104-155, 171-198



FIGURE 285

CGGACGCGTGGGCGGACGACGCTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCAGGTCCAGGTTTTGCTTTGA  
TCCTTTTCAAAAAGTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAACTACCCCT  
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCCAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC  
TCGGGAGTCGCTGCTTCCAAAGTGCCCGCGGTGAGTGAGCTCTCACCCCAGTGCAGCCAAATGAGCCTCTTCGGGC  
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC  
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG  
GAAGTATTCACAGCCCAAGGTTTTCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG  
AGGAAAATGTATGGATACAACCTTACGTTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT  
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACCTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG  
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCTTCTGAACCAGGGT  
TCTGCATCCACTACAACATTGTTCATGCCACAATTCACAGAAGCTGTGAGTCCTTCAGTGCCTACCCCTTCAGCTT  
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGGGAAGACCTTATTCGATATCTTGAACCAG  
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAAGTGGCAACTTCTTGGCAAGGCTTTTGTGTTTTGGAA  
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACCTCT  
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG  
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC  
ACGAGGTCTTTCAGTTGAGACCAAAGACCGGTGTGAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGC  
ACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACACCAGCAGCTCTTGCCCA  
GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTTGCT  
TCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA  
ACAGCTCTTTTGGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAAGAAAATTAATGTTGTAT  
TAAATAGATCACAGCTAGTTTTCAGAGTTACCATGTACGTATTCCTAGCTGGGTTCTGTATTTTCAGTTCTTTC  
GATACGGCTTAGGGTAATGTGAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAAC  
TCTAAAGCTCCATGTCTTGGGCCCTAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTTTGGCTCATATTCACAT  
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAACCTGTG  
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAATAATTTCTGCCATTTAGAAGAAGAGAACTACA  
TTCATGGTTTGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG  
TTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCT  
ATTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCT  
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA  
TTCTCGTATGGTGTAGAGTTAGATTAATCTGCATTTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA  
GACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA  
AAGTAGACATTTCAGATCCAGCCATTACTAACCTATTCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT  
AAAGCACCTTGAAAAGACTTGGCAGCTTCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTA  
TTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAATACATGGATATTTTTATGTACA  
GAAGTATGTCTCTTAACCAGTTCACCTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAAATATTTTGCTTGT  
AAAATGCTTAATATNGTGCCTAGGTATGTGGTGAATTTTGAATCAAAAATGTATTGAATCATCAATAAAAAGA  
ATGTGGCTATTTTGGGGAGAAAATTAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

**FIGURE 286**

MSLFGLLLLT SALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPRF  
PHTYPRNTVLVWRLVAVEENVWIQLTDFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWCGS  
GTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLLNNA  
ITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLEEVRLYSCTP  
RNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQLRPKT  
GVRGLHKSLTDVALEHHEECDVCVCRGSTGG

**Important features:****signal sequence:**

Amino acids 1-14

**N-glycosylation sites:**

Amino acids 25-29;55-59;254-258

**N-myristoylation sites:**

Amino acids 15-21;117-123;127-133;281-287;282-288;319-325

**Amidation site:**

Amino acids 229-233

**FIGURE 287**

CAGCGCTGACTGCGCCGCGGAGAAAGCCAGTGGGAACCCAGACCCATAGGAGACCCGCGTCCC  
CGCTCGGCCTGGCCAGGCCCCGCGCTATGGAGTTCTCTCTGGGCCCCCTCTCTTGGGTCTGTGCT  
GCAGTCTGGCCGCTGCTGATCGCCACACCGTCTTCTGGAACAGTTCAAATCCCAAGTTCCGGA  
ATGAGGACTACACCATAACATGTGCAGCTGAATGACTACGTGGACATCATCTGTCCGCACTATG  
AAGATCACTCTGTGGCAGACGCTGCCATGGAGCAGTACATACTGTACCTGGTGGAGCATGAGG  
AGTACCAGCTGTGCCAGCCCCAGTCCAAGGACCAAGTCCGCTGGCAGTGAACCGGCCCCAGTG  
CCAAGCATGGCCCCGGAGAAGCTGTCTGAGAAGTTCCAGCGCTTCACACCTTTCACCCTGGGCA  
AGGAGTTCAAAGAAGGACACAGCTACTACTACATCTCCAAACCCATCCACCAGCATGAAGACC  
GCTGCTTGAGGTTGAAGGTGACTGTCAGTGGCAAAATCACTCACAGTCCTCAGGCCCCATGACA  
ATCCACAGGAGAAGAGACTTGCAGCAGATGACCCAGAGGTGCGGGTTCTACATAGCATCGGTC  
ACAGTGCTGCCCCACGCCTCTTCCCACCTTGCTGGACTGTGCTGCTCCTTCCACTTCTGCTGC  
TGCAAACCCCGTGAAGGTGTGTGCCACACCTGGCCTTAAAGAGGGACAGGCTGAAGAGAGGGA  
CAGGCACTCCAAACCTGTCTTGGGGCCACTTTCAGAGCCCCCAGCCCTGGGAACCACTCCAC  
CACAGGCATAAGCTATCACCTAGCAGCCTCAAACGGGTCAATATTAAGGTTTTCAACCGGAA  
GGAGGCCAACCAGCCCGACAGTGCCATCCCCACCTTCACCTCGGAGGGATGGAGAAAGAAGTG  
GAGACAGTCCTTTCCACCATTCCTGCCTTTAAGCCAAAGAAACAAGCTGTGCAGGCATGGTC  
CCTTAAGGCACAGTGGGAGCTGAGCTGGAAGGGGGCCACGTGGATGGGCAAAGCTTGTCAAAGA  
TGCCCCCTTCAGGAGAGAGCCAGGATGCCCAGATGAACTGACTGAAGGAAAAGCAAGAAACAG  
TTTCTTGCTTGGAAGCCAGGTACAGGAGAGGCAGCATGCTTGGGCTGACCCAGCATCTCCCAG  
CAAGACCTCATCTGTGGAGCTGCCACAGAGAAGTTTGTAGCCAGGTACTGCATTCTCTCCCAT  
CCTGGGGCAGCACTCCCCAGAGCTGTGCCAGCAGGGGGGCTGTGCCAACCTGTTCTTAGAGTG  
TAGCTGTAAGGGCAGTGCCCATGTGTACATTCTGCCTAGAGTGTAGCCTAAAGGGCAGGGCCC  
ACGTGTATAGTATCTGTATATAAGTTGCTGTGTGTCTGTCCTGATTTCTACAACCTGGAGTTTT  
TTTATACAATGTTCTTTGTCTCAAATAAAGCAATGTGTTTTTTTCGG

FIGURE 288

MEFLWAPLLGLCCSLAAADRHTVFWNSSNPKFRNEDYTIHVQLNDYVDIICPHYEDHSADAAM  
EQYILYLVEHEEYQLCQPQSKDQVRWQCNRPSAKHGPEKLSEKFQRFPTFTLGKEFKEGHSYY  
YISKPIHQHEDRCLRLKVTVSGKITHSPQAHDNPQEKRLAADDPEVRVLHSIGHSAAPRLFPL  
AWTVLLLPLLLLQTP

Important features:

Signal sequence:

Amino acids 1-17

N-glycosylation site:

Amino acids 26-30

Tyrosine kinase phosphorylation site:

Amino acids 118-127

N-myristoylation site:

Amino acids 10-16

**FIGURE 289**

CGGACGCGTGGGCGGACGCGTGGGCGGCCCCACGGCGCCCGCGGGCTGGGGCGGTTCGCTTCTTC  
CTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAG**ATG**GCCCCATGGCCCCGAAGGGC  
CTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGGCTC  
CAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTGCCGG  
GGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACCTTTGGAGGTGGA  
AACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTGGTAGAG  
GTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCTGAGTGAG  
GAGCTGGTGGAGAGCTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCCAGTGGCTG  
TGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGGCCCTCCTGCCTTCCCTGT  
CCTGGGGGAACAGAGAGGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGGGACACGAGGG  
GGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCCAGTGTGGCCTT  
GGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGTTTTGGCCCCCTGT  
GCCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGCAATGCAAGAAGGGCTGGGCCCTGCAT  
CACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACTGTGGAGCTGACCAA  
TTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAGGCCTGCCTAGGCTGC  
ATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCAGCAGGTGGGCTCCAAG  
TGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGAACAAGCAGTGTGAAAAC  
ACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAGATGGAAGGCATCTGTGTG  
AAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGACAGAAGACGAGTTGGTGGTG  
CTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCACGCTGGCTGCTAAGGGCGAC  
TTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATGACTGGCTACTGGTTGTCAGAG  
CGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGAT**TAA**TCGCGGCCACCACCTGTAGGA  
CCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCCTCCTGCTGGACACTCAGGACAGC  
TTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTGCCTTACAGAGCAGCCCAGGTACCC  
AGGCCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGCCCTGAAGGTGGATACCATGAGCTCT  
TCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGAATTTCAAAGTTTTTCCTTAATGGTG  
GCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTGGTCCTCACAGGGGTGGGGCCATCACAG  
CTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGTTCTGTGTTACCCACATCCCCACACCCCA  
TTGCCACTTATTTATTCATCTCAGGAAATAAAGAAAGGTCTTGGAAGTTAAAAAAAAAAAAA  
AAAAAAAAAAAA

## **FIGURE 290**

MAPWPPKGLVPAVLWGLSLFLNLP GPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLERT  
 IRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQQE  
 APDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAGYGG  
 EACGQCGLGYFEAERNASHLVCSACFGPCARCSGPPEESNCLQCKKGWALHHLKCVDIDECGTE  
 GANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECETEVCP  
 GENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVV LQQMFFGIIICAL  
 ATLAAGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

### **Important features:**

#### **Signal sequence:**

Amino acids 1-29

#### **Transmembrane domain:**

Amino acids 342-392

#### **N-glycosylation sites:**

Amino acids 79-83;205-209

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 290-294

#### **Aspartic acid and asparagine hydroxylation site:**

Amino acids 321-333

#### **EGF-like domain cysteine pattern signature:**

Amino acids 181-193

**FIGURE 291**

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCCCTCGACCTCGACCCAC  
 GCGTCCGAACACAGGTCCTTGTTGCTGCAGAGAAGCAGTTGTTTTGCTGGAAGGAGGGAGTGCGCGGGCTGCCCC  
 GGGCTCCTCCCTGCCGCCTCCTCTCAGTGGATGGTTCCAGGCACCCTGTCTGGGGCAGGGAGGGACAGGCCTGC  
 ACATCGAAGGTGGGGTGGGACCAGGCTGCCCCCTCGCCCCAGCATCCAAGTCTCCCTTGGGCGCCCCGTGGCCCTG  
 CAGACTCTCAGGGCTAAGGTCCTCTGTTGCTTTTTGTTCCACCTTAGAAGAGGCTCCGCTTGACTAAGAGTAGC  
 TTGAAGGAGGCACCATGCAGGAGCTGCATCTGCTCTGGTGGGCGCTTCTCCTGGGCCTGGCTCAGGCCTGCCCTG  
 AGCCCTGCGACTGTGGGGAAAAGTATGGCTTCCAGATCGCCGACTGTGCCTACCGCGACCTAGAATCCGTGCCGC  
 CTGGCTTCCCGGCAATGTGACTACACTGAGCCTGTGAGCAACCGGCTGCCAGGCTTGCCGGAGGGTGCCTTCA  
 GGGAGGTGCCCCCTGCTGCAGTCGCTGTGGCTGGCACACAATGAGATCCGCACGGTGGCCGCCGGAGCCCTGGCCT  
 CTCTGAGCCATCTCAAGAGCCTGGACCTCAGCCACAATCTCATCTCTGACTTTGCCTGGAGCGACCTGCACAACC  
 TCAGTGCCCTCCAATTGCTCAAGATGGACAGCAACGAGCTGACCTTCATCCCCCGGACGCCTTCCGCAGCCTCC  
 GTGCTCTGCGCTCGCTGCAACTCAACCACAACCGCTTGCACACATTGGCCGAGGGCACCTTACCCCGCTCACCG  
 CGCTGTCCACCTGCAGATCAACGAGAACCCCTTCGACTGCACCTGCGGCATCGTGTGGCTCAAGACATGGGCCC  
 TGACCACGGCCGTGTCCATCCCGGAGCAGGACAACATCGCCTGCACCTCACCCCATGTGCTCAAGGGTACACCGC  
 TGAGCCGCCTGCCGCCACTGCCATGCTCGGCGCCCTCAGTGCAGCTCAGCTACCAACCCAGCCAGGATGGTGCCG  
 AGCTGCGGCCTGGTTTTGTGCTGGCACTGCACTGTGATGTGGACGGGCAGCCGGCCCCCTCAGCTTCACTGGCACA  
 TCCAGATACCCAGTGGCATTGTGGAGATCACCAGCCCCAACGTGGGCACTGATGGGCGTGCCTTGCCTGGCACCC  
 CTGTGGCCAGCTCCCAGCCGCGCTTCCAGGCCTTTGCCAATGGCAGCCTGCTTATCCCCGACTTTGGCAAGCTGG  
 AGGAAGGCACCTACAGCTGCCTGGCCACCAATGAGCTGGGCAGTGCTGAGAGCTCAGTGGACGTGGCACTGGCCA  
 CGCCCGGTGAGGGTGGTGGAGACACACTGGGGCGCAGGTTCCATGGCAAAGCGGTTGAGGGAAAGGGCTGCTATA  
 CGGTTGACAACGAGGTGCAGCCATCAGGGCCGGAGGACAATGTGGTCATCATCTACCTCAGCCGTGCTGGGAACC  
 CTGAGGCTGCAGTCGAGAAGGGTCCCTGGGCAGCTGCCCCCAGGCCTGCTCCTGCTGGGCCAAAGCCTCCTCC  
 TCTTCTTCTTCTCCTCACCTCCTTCTAGCCCCACCCAGGGCTTCCCTAACTCCTCCCCCTTGCCCCCTACCAATGCCCC  
 TTTAAGTGCTGCAGGGGTCTGGGGTTGGCAACTCCTGAGGCCTGCATGGGTGACTTACATTTTCTACCTCTCC  
 TTCTAATCTCTTCTAGAGCACCTGCTATCCCCAACTTCTAGACCTGCTCCAACTAGTGAAGTAGGATAGAATTTG  
 ATCCCCTAACTCACTGTCTGCGGTGCTCATTGCTGCTAACAGCATTGCCTGTGCTCTCCTCTCAGGGGCAGCATG  
 CTAACGGGGCGACGTCTAATCCAACCTGGGAGAAGCCTCAGTGGTGGAATTCCAGGCACTGTGACTGTCAAGCTG  
 GCAAGGGCCAGGATTGGGGGAATGGAGCTGGGGCTTAGCTGGGAGGTGGTCTGAAGCAGACAGGGAATGGGAGAG  
 GAGGATGGGAAGTAGACAGTGGCTGGTATGGCTCTGAGGCTCCCTGGGGCCTGCTCAAGCTCCTCCTGCTCCTTG  
 CTGTTTTCTGATGATTTGGGGGCTTGGGAGTCCCTTTGTCTCATCTGAGACTGAAATGTGGGGATCCAGGATGG  
 CCTTCTTCTTCTTACCCTTCTCCCTCAGCCTGCAACCTCTATCCTGGAACCTGTCTCCCTTTCTCCCCAACT  
 ATGCATCTGTTGTCTGCTCCTCTGCAAAGGCCAGCCAGCTTGGGAGCAGCAGAGAAATAAACAGCATTCTGATG  
 CCAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCT

**FIGURE 292**

MQELHLLWWALLLGLAQACPEPCDCGEKYGFQIADCAYRDLESVPPGF PANVTTL SLSANRLP  
 GLPEGAFREVPLLQSLWLAHNEIRTVAAGALASLSHLKSLDL SHNLISDFAWSDLHNL SALQL  
 LKMDSNELTFIPRDAFRSLRALRSLQLNHNRLHTLAEGTFTPLTALSHLQINENPFDCTCGIV  
 WLKTWALTAVSIPEQDNIACTSPHVLKGTPLSRLPPLPCSAPSVQLSYQPSQDGAELRPGFV  
 LALHCDVDGQPAPQLHWHIQIPSGIVEITSPNVGTDGRALPGTPVASSQPRFQAFANGSLLIP  
 DFGKLEEGTYSCLATNELGSAESSVDVALATPGE GGEDTLGRRFHGKAVEGKGCYTV DNEVQP  
 SGPEDNVVVIYLSRAGNPEAAVAEGVPGQLPPGLLLL GQSLLLFFFLTSF

**Important features:****Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 403-418

**N-glycosylation sites:**

Amino acids 51-55,120-124,309-313

**Tyrosine kinase phosphorylation site:**

amino acids 319-326

**N-myristoylation sites:**amino acids 14-20,64-70,92-98,218-224,294-300,323-329,334-340,  
350-356,394-400**Amidation site:**

amino acids 355-359

**Leucine Rich Repeat:**

amino acids 51-74,75-98, 99-122,123-146,147-170

**Leucine rich repeat C-terminal domain:**

amino acids 180-230





**FIGURE 294**

MEKMLAGCFLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFIIID  
 SSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAVKR  
 MRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARDTGI  
 LIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHNCAHF  
 CINIPGSYVCRCKQGYILNSDQTTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYALAEDEG  
 KRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGEHECVN  
 MEESYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLLINEDLKTC  
 SRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSALGDHGCEHSCVSSSED  
 SFVCQCFCFEGYILREDGKTCRRKDVCAIDHGCEHICVNSDDSYTCECLEGFRLAEDGKRCRRK  
 DVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVIDGSKSLGEENF  
 EVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVAHMKYMGKGSMT  
 GLALKHMFERSFTQEGEARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKANGITMYAVGVGKA  
 IEELQEIASPTNKHLFYAEDFSTMDIEISEKLKKGICEALEDSQDGRQDSPAGELPKTVQQPT  
 ESEPVITINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPLEEKHDQCKCENLIM  
 FQNLANEEVRKLTQRLEEMTQRMALLENRLRYR

**Important features:****Signal sequence:**

Amino acids 1-23

**N-glycosylation site:**

Amino acids 221-225

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 115-119;606-610;892-896

**N-myristoylation sites:**

Amino acids 133-139;258-264;299-305;340-346;453-459;494-500;  
 639-645;690-694;  
 752-758;792-798

**Amidation sites:**

Amino acids 314-318;560-564;601-605

**Aspartic acid and asparagine hydroxylation sites:**

Amino acids 253-265;294-306;335-347;376-388;417-429;  
 458-470;540-552;581-593

GGCCGGAGCAGCACGGCCGACGGACCTGGAGCTCCGGCTGCGTCTTCCCGCAGCGCTACCCGC  
C**ATG**CGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTGCTGCTGCCGCCCGC  
GCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCTGGTGGACAAGTTTAA  
CCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACACGGCTTGGGAGGAAAA  
GACGCTGTCCAAGTACGAGTCCAGCGAGATTGCGCTGCTGGAGATCCTGGAGGGGCTGTGCGA  
GAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGAGCACCTGGAGGCCTGGTG  
GCTGCAGCTGAAGAGCGAATATCCTGACTTATTTCGAGTGGTTTTGTGTGAAGACACTGAAAGT  
GTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGCCAGGGCGGATCCCAGAGGCC  
CTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACAGGGCGACGGGTCTGCCGGTG  
CCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGGACGGCTACTTCAGCTCGCTCCG  
GAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCCTGCAAGACGTGCTCGGGCCTGAC  
CAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCTGGACGAGGGCGCCTGTGTGGATGT  
GGACGAGTGTGCGGCCGAGCCGCCCTCCCTGCAGCGCTGCGCAGTTCTGTAAGAACGCCAACGG  
CTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTGGGCTGCACAGGGGAAGGCCCAGGAAA  
CTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCACGGACAGTGTGCAGATGTGGACGAGTG  
CTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACGAAAACCTGCTACAATACTCCAGGGAGCTA  
CGTCTGTGTGTGTCCTGACGGCTTCGAAGAAACGGAAGATGCCTGTGTGCCGCCGGCAGAGGC  
TGAAGCCACAGAAGGAGAAAGCCCGACACAGCTGCCCTCCCGCGAAGACCTG**TAA**TGTGCCGG  
ACTTACCCTTTAAATTTATTTCAGAAGGATGTCCCGTGGAATGTGGCCCTGAGGATGCCGTCT  
CCTGCAGTGGACAGCGGCGGGGAGAGGCTGCCTGCTCTCTAACGGTTGATTCTCATTTGTCCC  
TTAAACAGCTGCATTTCTTGGTTGTTCTTAAACAGACTTGTATATTTTGATACAGTTCTTTGT  
AATAAAATTGACCATTGTAGGTAATCAGGAGGAAAAA

## **FIGURE 296**

MRLPRRAALGLLPLLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEK  
 TLSKYESSEIRLLEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTALKV  
 CCSPGTYGPDCLACQGGSQRPCSGNGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLR  
 NETHSICTACDESKTCSGLTNRDCGECEVGWVLDEGACVDVDECAAEPFPCSAQFCKNANG  
 SYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKTCVRKNENCYNTPGSY  
 VCVCPDGFEEETEDACVPPAEAEATEGESPTQLPSREDL

### **Important features:**

#### **Signal peptide:**

Amino acids 1-24

#### **N-glycosylation sites:**

Amino acids 190-194;251-255

#### **Glycosaminoglycan attachment sites:**

Amino acids 149-153;155-159

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 26-30

#### **Tyrosine kinase phosphorylation site:**

Amino acids 303-310

#### **N-myristoylation sites:**

Amino acids 44-50;54-60;55-61;81-87;150-156;158-164;164-170;  
 252-258;313-319

#### **Aspartic acid and asparagine hydroxylation site:**

Amino acids 308-320

#### **EGF-like domain cysteine pattern signature:**

Amino acids 166-178

#### **Leucine zipper pattern:**

Amino acids 94-116

296/550



**FIGURE 298**

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLTF  
 PARLPANTQIILLQTNNAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLEEN  
 KLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDALPN  
 LEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRLIKVP  
 HVALQKVVNLFKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLPDLRKIE  
 ATNNPRLSYIHPNAFFRLPKLESIMLNSNALSALYHGTIESLPNLKEISIHSPNPIRCDCVIRW  
 MNMNKTNIRFMEPDSLFCVDPPEFQGQNVVRQVHFRDMMEICLPLIAPESFSPSNLNVEAGSYVS  
 FHCRATAEPQPEIYWITPSGQKLLPNTLTDFYVHSEGTLDINGVTPKEGGLYTCIATNLVGA  
 DLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFVKTENSAAQ  
 SARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKEYEKNNTTTLM  
 ACLGGLLGIIGVICLISCLSPENMCDGGHSYVRNYLQKPTFALGELYPPLINLWEAGKEKSTS  
 LKVKATVIGLPTNMS

**Important features:****Signal sequence:**

amino acids 1-22

**Transmembrane domain:**

amino acids 633-650

**N-glycosylation site.**amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,  
608-612, 624-628, 625-629**Casein kinase II phosphorylation site.**

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

**Tyrosine kinase phosphorylation site.**

amino acids 570-579

**N-myristoylation site.**amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,  
491-497, 492-498, 634-640, 702-708**Cell attachment sequence.**

amino acids 277-280

**FIGURE 299**

GCTGTGGGAACCTCTCCACGCGCACGAACTCAGCCAACGATTTCTGATAGATTTTTGGGAGTT  
TGACCAGAGATGCAAGGGGTGAAGGAGCGCTTCCTACCGTTAGGGAACTCTGGGGACAGAGCG  
CCCCGGCCGCCTGATGGCCGAGGCAGGGTGCGACCCAGGACCCAGGACGGCGTCGGGAACCAT  
ACC**ATG**GGCCGGATCCCCAAGACCCCTAAAGTTCGTCGTCGTCATCGTCGCGGTCCTGCTGCCA  
GTCCTAGCTTACTCTGCCACCACTGCCCCGGCAGGAGGAAGTTCCCCAGCAGACAGTGGCCCCA  
CAGCAACAGAGGCACAGCTTCAAGGGGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAACAT  
ACTGGAGCCTGTAACCCGTGCACAGAGGGTGTGGATTACACCAACGCTTCCAACAATGAACCT  
TCTTGCTTCCCATGTACAGTTTGTAAATCAGATCAAAAACATAAAAGTTCCTGCACCATGACC  
AGAGACACAGTGTGTCAGTGTAAAGAAGGCACCTTCCGGAATGAAAACCTCCCCAGAGATGTGC  
CGGAAGTGTAGCAGGTGCCCTAGTGGGGAAGTCCAAGTCAGTAATTGTACGTCCTGGGATGAT  
ATCCAGTGTGTTGAAGAATTTGGTGCCAATGCCACTGTGGAAACCCCAGCTGCTGAAGAGACA  
ATGAACACCAGCCCGGGGACTCCTGCCCCAGCTGCTGAAGAGACAATGAACACCAGCCCAGGG  
ACTCCTGCCCCAGCTGCTGAAGAGACAATGACCACCAGCCCGGGGACTCCTGCCCCAGCTGCT  
GAAGAGACAATGACCACCAGCCCGGGGACTCCTGCCCCAGCTGCTGAAGAGACAATGACCACC  
AGCCCGGGGACTCCTGCCTCTTCTCATTACCTCTCATGCACCATCGTAGGGATCATAGTTCTA  
ATTGTGCTTCTGATTGTGTTTGT**TGA**AAGACTTCACTGTGGAAGAAATTCCTTCCTTACCTG  
AAAGGTTCAGGTAGGCGCTGGCTGAGGGCGGGGGCGCTGGACACTCTCTGCCCTGCCTCCCT  
CTGCTGTGTTCCACAGACAGAAACGCCTGC

**FIGURE 300**

MARIPKTLKFVVVIVAVLLPVLAYSATTARQEEVPQQTVA PQQRHSFKGEECPAGSHRSEHT  
GACNPCTEGVDYTNASNNEPSCFPCTVCKSDQKHKSCTMTRDTVCQCKEGTFRNENSPEMCR  
KCSRCPSGEVQVSNCTSWDDIQCVEEFGANATVETPAAEETMNTSPGTPAPAAEETMNTSPGT  
PAPAAEETMTTSPGTPAPAAEETMTTSPGTPAPAAEETMTTSPGTPASSHYLSCTIVGIIVLI  
VLLIVFV

**Important features:****Signal peptide:**

Amino acids 1-29

**Transmembrane domain:**

Amino acids 240-259

**N-glycosylation site:**

Amino acids 77-81;140-144;156-160

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 126-130

**N-myristoylation sites:**

Amino acids 56-62;72-78;114-120;154-160;233-239



**FIGURE 301**

CACAAGCATCTTAATTTGAATCCACAAAGTTTCATGTAATGAAAAGAAATACATAATTTTAAT  
 TCAACCCGAGTGTTTTCCAAGAAGATTGTATTTGCTTAAATTGCTACAGTAATTCAAGAGACA  
 GCCCTGTCTGGACACAGAGTTACTGTGGATTTTTAAGAGACTCAGTTAAAGAATTTAGGAATT  
 TCTGATTCATTTAAAGGATTTACAAATTCATCAACCCCTGAAAACCTAAAGCAAATTGAACAGG  
 AAAAAAAAAAAGAAG**ATG**GGTTTTTTAAGTCCAATATATGTTATTTTCTTCTTTTTTGGAGTC  
 AAAGTACATTGCCAATATGAACTTATCAGTGGGATGAAGACTATGACCAAGAGCCAGATGAT  
 GATTACCAAACAGGATTCCCATTTTCGTCAAATGTAGACTACGGAGTTCCTTTTCATCAGTAT  
 ACTTTAGGCTGTGTCTCAGTGAATGCTTCTGTCCAACCTAAGTTTCCATCATCAATGTACTGTGAT  
 AATCGCAAACCTCAAGACTATCCCAAATATTCGGATGCACATTCAGCAACTCTACCTTCAGTTC  
 AATGAAATTGAGGCTGTGACTGCAAATTCATTCATCAATGCAACTCATCTTAAAGAAATTAAC  
 CTCAGCCACAACAAAATTAAATCTCAAAAGATTGATTATGGTGTGTTTGCTAAGCTTCCAAAT  
 CTACTACAACCTTCATCTAGAGCATAATAATTTAGAAGAATTTCCATTTCTCCTTAAATCT  
 CTGGAAAGACTCCTTCTTGGTTACAATGAAATCTCCAACTGCAGACAAATGCTATGGATGGG  
 CTAGTAAACTTGACCATGCTTGATCTCTGTTATAATTATCTTCATGATTCTCTGCTAAAAGAC  
 AAAATCTTTGCCAAAATGGAAAACTAATGCAGCTCAACCTCTGCAGTAACAGATTAGAATCA  
 ATGCCTCCTGGTTTGCCTTCTTCACTTATGTATCTGTCTTTAGAAAATAATTCAATTTCTTCT  
 ATACCCGAAAAATACTTCGACAACTTCCAAACTTCATACTCTAAGAATGTCACACAACAAA  
 CTACAAGACATCCCATATAATATTTTTAATCTTCCCAACATTGTAGAACTCAGTGTTGGACAC  
 AACAAATTGAAGCAAGCATCTATATTCCAAGAAATTTGGAACACCTATACCTACAAAATAAT  
 GAAATAGAAAAGATGAATCTTACAGTGATGTGTCTTCTATTGACCCACTACATTACCACCAT  
 TTAACATACATTCGTGTGGACCAAAATAAACTAAAAGAACCAATAAGCTCATAACATCTTCTTC  
 TGCTTCCCTCATATACACACTATTTATTATGGTGAACAACGAAGCACTAATGGTCAAACAATA  
 CAACTAAAGACACAAGTTTTTCAGGAGATTTCCAGATGATGATGATGAAAGTGAAGATCACGAT  
 GATCCTGACAATGCTCATGAGAGCCCAGAACAAGAAGGAGCAGAAGGGCACTTTGACCTTCAT  
 TATTATGAAAATCAAGAAT**TAG**CAAGAACTATATAGGTATACACTTACGACTTCACAAAACCTA  
 TACTTAATATAGTAAATCTAAGTAAACATGTATTACTCAAAGTAATATATTTAGAATTATGTA  
 TTAGTATAAGATCAGAATTGAATTTAAGTTGTTGGTGACATCTGCATCATTTTCATAGGATTAG  
 AACTTACTCAAATAATGTAAATCTTTAAAAATATAAATTAGAATGACAAGTGGGAATCATAA  
 ATTAAACGTTAATGGTTTCTTATGCTCTTTTTAAATATAGAAATATCATGTTAAAGAAAAAA  
 AAAAAA

**FIGURE 302**

MGFLSPIYVIFFFFGVKVHCQYETYQWDEDYDQEPDDDYQTGFPPFRQNVVDYGVPFHQYTLGCV  
SECFCPTNFPSSMYCDNRKLKTIPNIPMHIQQLYLQFNEIEAVTANSFINATHLKEINLSHNK  
IKSQKIDYGVFAKLPNLLQLHLEHNNLEEFPPPLPKSLERLLLGYNEISKLQTNAMDGLVNLT  
MLDLCYNYLHDSLLKDKIFAKMEKLMQLNLCSNRLESMPGLPSSLMYLSLENNSSIPIPEKY  
FDKLPKLHTLRMSHNKLQDIPYNIFFNLPNIVELSVGHNKLKQAFYIPRNLEHLYLQNNIEIEKM  
NLTVMCPSIDPLHYHHLTYIRVDQNKLEPISSYIFFCFPHIHTIYYGEQRSTNGQTIQLKTQ  
VFRFPDDDDDESEDHDDPDNAHESPEQEGAEGHFDLHYENQE

**Important features:****N-glycosylation sites:**

Amino acids 113-117;121-125; 187-191;242-246;316-320

**Tyrosine kinase phosphorylation sites:**

Amino acids 268-275;300-307

**N-myristoylation site:**

Amino acids 230-236

**Leucine zipper patterns:**

Amino acids 146-168;217-239

**FIGURE 303**

GCCCCGGACTGGCGCAAGGTGCCCAAGCAAGGAAGAAATAATGAAGAGACACATGTGTAGC  
TGCAGCCTTTTGAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTACCA  
CGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGC**ATGA**ATCTGGT  
AGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTGTTCCTTATGAT  
ACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGGGGTTT  
AAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGAAACAGT  
CTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGGACCTCCA  
TCAACTGAGAGTTCTCAACCTGTCCAAAATGGCATTGAGTTTATCGATGAGCATGCCTTCAA  
AGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAGTGTGCACAA  
AAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCACAACCCCTGGCACTGCGACTG  
TACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCACAACGTGATCTG  
TAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTTCCTCAATGCTGCCAACGACGCTGA  
CCTTTGTAACTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCATGTTTGGCTGGTT  
CACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAGGATGCCCGGAGACA  
CCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGAACCTGATGATATTAG  
CACTGTGGTA**TAG**TGTGCCAACTGACTGTCATTGAGAAAGAAAGAAAGTAGTTTGCATTGCA  
GTAGAAATAAGTGTTTACTTCTCCCATCCATTGTAAACATTTGAAACTTTGTATTTAGTTT  
TTTTTGAATTATGCCACTGCTGAACTTTTAAACAAACACTACAACATAAATAATTTGAGTTTAG  
GTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGTAAGCTACTATCTGAACATT  
AGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTTAATTTAAAAGCAAATAAAAG  
CTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAAAAAACA

**FIGURE 304**

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDLP  
PETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDLSDNRIQ  
SVHKNAFNNLKARARIANNPWHCCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFLNAA  
NDADLCNLPKKTTDYAMLVTMFGWFTMVISYVYYYVRQNQEDARRHLEYLKSLPSRQKKADEP  
DDISTVV

**Important features:****Signal sequence:**

Amino acids 1-33

**Transmembrane domain:**

Amino acids 204-219

**N-glycosylation sites:**

Amino acids 47-51;94-98

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 199-203

**Casein kinase II phosphorylation site.**

amino acids 162-166, 175-179

**N-myristoylation sites:**

Amino acids 37-43;45-51;110-116



**FIGURE 306**

MKRLPLLVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNECGNLTQSCGENANC  
 TNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENTVNANCHLDNVCIAANINKTLTKIRSIKEPVALLQEVYRNS  
 VTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTLSNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKL  
 MHTVEQATLRISQSFQKTTEFDTNSTDIALKVFFFDSDYNMKHIHPHMNDGDYINIFPKRKAAYDSNGNVAVAF  
 YYKSIGPLLSSSDNFLKPKQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDYRSLCAFWNY  
 SPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLGIIISLICLAICIFTFW  
 FFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSEIAGLLHYFFLAFAWMCIEGIHLYLIVGVIYN  
 KGFLHKNFYIFGYLSPAVVVGFSAAALGYRYYGTTKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVRHT  
 AGLKPEVSCFENIRSCARGALALLFLLGTTWIFGVLVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEY  
 YRLFKNVPCCFGCLR

**Important features:****Signal peptide:**

Amino acids 1-19

**Transmembrane domain:**

Amino acids 431-450;494-515;573-594;619-636;646-664

**N-glycosylation sites:**

Amino acids 15-19;21-25;64-68;74-78;127-131;177-181;  
 188-192;249-253;381-385;395-399

**Glycosaminoglycan attachment site:**

Amino acids 49-53

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 360-364

**Tyrosine kinase phosphorylation sites:**

Amino acids 36-44;670-677

**N-myristoylation sites:**

Amino acids 38-44;50-56;52-58;80-86;382-388;388-394;  
 434-440;480-486;521-527

**Aspartic acid and asparagine hydroxylation site:**

Amino acids 75-87



## **FIGURE 308**

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGKM  
APFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLLGT  
VPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGCRNG  
GFCNERRICECPDGFHGPHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCSTTCFN  
GGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIKSKCKCSKGYQGDLCSPVCEPGCG  
AHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERRDPPESNYIW

**Important features:**

**Signal sequence:**

Amino acids 1-28

**N-glycosylation sites:**

Amino acids 88-92;245-249

**Tyrosine kinase phosphorylation site:**

Amino acids 370-378

**N-myristoylation sites:**

Amino acids 184-190;185-191;189-195;315-321

**ATP/GTP-binding site motif A (P-loop):**

Amino acids 285-293

**EGF-like domain cysteine pattern signatures:**

Amino acids 198-210;230-242;262-274;294-306;326-338



CCACACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGGG  
TTAGACTGGCGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGACT  
CTTGCAAGCTGGATGCCCTCTGTGGATGAAAG**ATG**TATCATGGAATGAACCCGAGCAATGGAG  
ATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTGGCCG  
TGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTTTCGATC  
ACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCGGAGGGG  
TTTTCTTTGAAGGCTCTGTAGCCCGATTTCACCTGCCAAGACGGATTCAAGCTGAAGGGCGCTA  
CAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGATAATTCCA  
TCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCTATAACAAGACAT  
ATAGACATGGAGAGAAGCTAATCATCACTTGTTCATGAAGGATTCAAGATCCGGTACCCCGACC  
TACACAATATGGTTTCATTATGTTCGCATGATGGAACGTGGAATAATCTGCCCATCTGTCAAG  
GCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGCTCCAGACCTCCT  
TCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTTGATGGGTCTGCGT  
ATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTGCCTTGCTCTGGAAG  
CCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTTCGTCTGCCACCCGCGGCCTT  
GTGAGCGCTACAACCACGGAACGTGTGGTGGAGTTTTACTGCGATCCTGGCTACAGCCTCACCA  
GCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTTCTTCTTATCAAGTCTACTGCA  
TCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGACCACGTGGAAGATTGTGG  
CGTTCACGGACAACCAAGTGTGCTGCTGGTGTGCTGCTGCTGCTCATCCTGGCCAGGATGTTCCAGA  
CCAAGTTCAAGGCCCACTTTCCCCCAGGGGGCCTCCCCGGAGTTCCAGCAGTGACCCTGACT  
TTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATGACGAAGCTGTGAGTGGCGGCT  
TGAGTGCCTTAGGCCCCCGGTACATGGCCTCTGTGGGCCAGGGCTGCCCCTTACCCGTGGACG  
ACCAGAGCCCCCAGCATAACCCGGCTCAGGGGACACGGACACAGGCCCAGGGGAGTCAGAAA  
CCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAAGTCTGTATTACCTCCCAGGTGCC  
AAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATAATTGCCAGCACGGCAGAGGAGGTGG  
CATCCACCAGCCCAGGCATCCATCATGCCCCTGGGTGTTGTTCCCTAAGAAAC**TGA**TTTGATTA  
AAAAATTTCCCAAGTGTCTGAAGTGTCTCTTCAAATACATGTTGATCTGTGGAGTTGATTCT  
CTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAAAGCTCTGATCCTTAAATTGCTATGCTG  
ATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTCCTGTTTCTTCTTGACACAGACTGATTA  
AATTAAGNAAAAA

**FIGURE 310**

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGIP  
 ENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRIPQ  
 IEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLASSNG  
 YVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLPPMVS  
 HGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQTWPSTH  
 ETLLTTWKIVAFTATSVLLVLLLVLARMFQTKFKAHFPPRGPPRSSSSDPDFVVVDGVPVML  
 PSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGD TDTGPGESETCDSVSGSSEL  
 LQSLYSPPRCQESTHPASDNPDI IASTAEVASTSPGIHHAHWVFLRN

**Important features:****Signal sequence:**

amino acids 1-41

**Transmembrane domain:**

amino acids 325-344

**N-glycosylation site.**

amino acids 104-108, 134-138, 192-196

**Casein kinase II phosphorylation site.**

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,  
 364-368, 380-384, 467-471, 468-472

**N-myristoylation site.**

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,  
 478-484

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 394-405

**FIGURE 311**

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGAG  
CGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCTGG  
AGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCGAGGCCCCAGCTCAGGCT  
CGTGGCCACCCACCAAGTTCCAGTGCCGCGACAGTGCGTTATGCGTGCCCCTCACCTGGCGCT  
GCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCCATGTA  
CCCAGAAAGGGCAATGCCACCGCCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCAGTGACT  
GCTCTGGGGGAACTGACAAGAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCAGGCGAGC  
TCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCACCCAGACT  
GTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAGGGGATGCCA  
CAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTACCTCTCTCAGGAATGCCACAACCATGG  
GGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTGCGGAATGCCACATCCTCCTCTGCCGGAG  
ACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGCTCAGTGCAAGCC  
TGGTCACCGCCACCCCTCCTCCTTTTGTCTGGCTCCGAGCCCAGGAGCGCCTCCGCCCCACTGG  
GGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGACCTCGCTGCCCTGAG  
GACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACAGGAGGAGAGCAGTGA  
TGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTTCTGGCCACGTGGAACC  
TCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCCTGGACACTCCCTATGGA  
GATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAGGGGCTGGCCCCAGGCAGC  
TCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTGCCCCGTCTGAGGGTGGCGA  
TTAAAGTTGCTTC

**FIGURE 312**

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCRT  
SGLCVPLTWRCRDRLDCSDGSDEEEECRIEPTQKGQCPPPPGLPCPCTGVSDCSGGTDKKLRN  
CSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVTLES  
VTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLLLLSW  
LRAQERLRPLGLLVAMKESLLLSEQKTSLP

**Important features:****Signal sequence:**

Amino acids 1-30

**Transmembrane domain:**

Amino acids 231-248

**N-glycosylation sites:**

Amino acids 126-130;195-199;213-217

**Casein kinase II phosphorylation site.**

amino acids 84-88, 140-144, 161-165, 218-222

**N-myristoylation sites:**

Amino acids 3-9;10-16;26-32;30-36;112-118;166-172;212-218;  
224-230;230-236;263-269

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 44-55

**Leucine zipper pattern:**

Amino acids 17-39



## **FIGURE 314**

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRADDCA  
LPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYPVLGTYWDNCNRCT  
CQENRQWHGGSRRHDQSHQPGQLWLAGWEPQRLLGHDPG

### **Important features:**

#### **N-glycosylation site.**

amino acids 78-82, 161-165

#### **Casein kinase II phosphorylation site.**

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,  
411-415

#### **N-myristoylation site.**

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,  
269-275, 378-384, 442-448

#### **Amidation site.**

amino acids 26-30, 318-322

#### **Eukaryotic thiol (cysteine) proteases histidine active site.**

amino acids 398-409

**FIGURE 315**

CGGACGCGTGGGCCCCTGGTGGGCCCAGCAAGATGGATCTACTGTGGATCCTGCCCTCCCTGT  
GGCTTCTCCTGCTTGGGGGGCCTGCCTGCCTGAAGACCCAGGAACACCCCAGCTGCCCAGGAC  
CCAGGGAAGTGAAGCCAGCAAAGTTGTCTCCTGCCAGTTGTCCCGGAGCTCCAGGAAGTC  
CTGGGGAGAAGGGAGCCCCAGGTCCTCAAGGGCCACCTGGACCACCAGGCAAGATGGGCCCCA  
AGGGTGAGCCAGGCCCCAGAACTGCCGGGAGCTGTTGAGCCAGGGCGCCACCTTGAGCGGCT  
GGTACCATCTGTGCCTACCTGAGGGCAGGGCCCTCCAGTCTTTTGTGACATGGACACCGAGG  
GGGGCGGCTGGCTGGTGTTCAGAGGCGCCAGGATGGTTCTGTGGATTTCTTCCGCTCTTGGT  
CCTCCTACAGAGCAGGTTTTGGGAACCAAGAGTCTGAATTCTGGCTGGGAAATGAGAATTTGC  
ACCAGCTTACTCTCCAGGGTAACTGGGAGCTGCGGGTAGAGCTGGAAGACTTTAATGGTAACC  
GTACTTTGCCCCACTATGCGACCTTCCGCCTCCTCGGTGAGGTAGACCACTACCAGCTGGCAC  
TGGGCAAGTTCTCAGAGGGCACTGCAGGGGATTCCCTGAGCCTCCACAGTGGGAGGCCCTTTA  
CCACCTATGACGCTGACCACGATTCAAGCAACAGCAACTGTGCAGTGATTGTCCACGGTGCCT  
GGTGGTATGCATCCTGTTACCGATCAAATCTCAATGGTCGCTATGCAGTGTCTGAGGCTGCCG  
CCCACAAATATGGCATTGACTGGGCCTCAGGCCGTGGTGTGGGCCACCCCTACCGCAGGGTTC  
GGATGATGCTTCGATAGGGGCACTCTGGCAGCCAGTGCCCTTATCTCTCCTGTACAGCTTCCGG  
ATCGTCAGCCACCTTGCCTTTGCCAACCACCTCTGCTTGCCTGTCCACATTTAAAAATAAAAT  
CATTTTAGCCCTTTCA

**FIGURE 316**

MDLLWILPSLWLLLLGGPACLKTQEHPSCPGPRELEASKVVLLPSCPGAPGSPGEKGAPGPQG  
PPGPPGKMGPKGEPGPRNCRELLSQGATLSGWYHLCLPEGRALPVFCDMDTEGGGWLVFQRRQ  
DGSVDFFRSWSSYRAGFGNQESEFWLGNENLHQLTLQGNWELRVELEDFNGNRTFAHYATFRL  
LGEVDHYQLALGKFSEGTAGDSLSLHSGRPFTTYDADHDSSNSNCAVIVHGAWWYASCYRSNL  
NGRYAVSEAAAHKYGIDWASGRGVGHPYRRVRMMLR

**Important features:****Signal peptide:**

Amino acids 1-16

**N-glycosylation site:**

Amino acids 178-182

**Glycosaminoglycan attachment site:**

Amino acids 272-276

**Tyrosine kinase phosphorylation site:**

Amino acids 188-197

**N-myristoylation sites:**

Amino acids 16-22;89-95;144-150;267-273

**Fibrinogen beta and gamma chains C-terminal domain signature:**

Amino acids 242-255





**FIGURE 318**

MLPCASCLPGSLLLWALLLLLLLGSASPQDSEEPDSYTECTDGYEWDPDSSQHCRDVNECLTIPE  
ACKGEMKCINHYGGYLCLPRSAAVINDLHGEGPPPPVPPAQHPNPCPPGYEPDDQDSCVDVDE  
CAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCQHRCVNLPGSFRCQCE  
PGFQLGPNNRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGFSCSDIDECYSSY  
LCQYRCVNEPGRFSCHCPQGYQLLATRLCQDIDECESGAHQCEAQTCVNFHGGYRCVDTNRC  
VEPYIQVSENRCLCPASNPLCREQPSSIVHRYMTITSERSVPADVFIQATSVYPGAYNAFQI  
RAGNSQGDIFYIRQINNVSAMLVLARPVGTGPREYVLDLEMVTMNSLMSYRASSVLRLTVFVGAYTF

**Important features:****Signal sequence:**

Amino acids 1-25

**N-glycosylation sites:**

Amino acids 198-202;394-398

**N-myristoylation sites:**

Amino acids 76-82;145-151;182-188;222-228;290-296;305-311;  
371-377;381-387

**Aspartic acid and asparagine hydroxylation sites:**

amino acids 140-152;177-189;217-229;258-270

**FIGURE 319**

GCTGGGGACATGAGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTC  
TCAAAGGTGCGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCTGGCCACCTCCCCGATGC  
CCGCTGGGAGTACCCCTGGTGCTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGGCTG  
GGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGCCAGCCCGGG  
GCAGGACCCGGTGGCCGGGGGGCCCTGTGCCTCTTGGCAGAGGACGACAGCAGCTGTGAGGTG  
AACGGCCGCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGC  
GAGGACGGCGGCTTCACCTGCGTGCCGCTGTGCAGCGAGGATGTGCGGCTGCCAGCTGGGAC  
TGCCCCCACCCCAGGAGGGTCGAGGTCCTGGGCAAGTGCTGCCCTGAGTGGGTGTGCGGCCAA  
GGAGGGGGACTGGGGACCCAGCCCCTTCCAGCCCCAAGGACCCAGTTTTCTGGCCTTGTCTCT  
TCCCTGCCCCCTGGTGTCCCCTGCCCAGAATGGAGCACGGCCTGGGGACCCTGCTCGACCACC  
TGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAGAACCGCTTCTGCCGACTGGAGACCCAG  
CGCCGCTGTGCCTGTCCAGGCCCTGCCACCCCTCCAGGGGTGCGAGTCCACAAAACAGTGCC  
TTCTAGAGCCGGGCTGGGAATGGGGACACGGTGTCCACCATCCCCAGCTGGTGGCCCTGTGCC  
TGGGCCCTGGGCTGATGGAAGATGGTCCGTGCCCAGGCCCTTGGCTGCAGGCAACACTTTAGC  
TTGGGTCCACCATGCAGAACACCAATATTAACACGCTGCCTGGTCTGTCTGGATCCCGAGGTA  
TGGCAGAGGTGCAAGACCTAGTCCCCTTTCCTCTAACTCACTGCCTAGGAGGCTGGCCAAGGT  
GTCCAGGGTCCTCTAGCCCACTCCCTGCCTACACACACAGCCTATATCAAACATGCACACGGG  
CGAGCTTTCTCTCCGACTTCCCCTGGGCAAGAGATGGGACAAGCAGTCCCTTAATATTGAGGC  
TGCAGCAGGTGCTGGGCTGGACTGGCCATTTTTCTGGGGGTAGGATGAAGAGAAGGCACACAG  
AGATTCTGGATCTCCTGCTGCCTTTTCTGGAGTTTGTAATAATTGTTCCCTGAATACAAGCCTAT  
GCGTGA

**FIGURE 320**

MRGTPKTHLLAFSLLCLLSKVRTQLCPTPCTCPWPPPRCPLGVPLVLDGCGCCRVCARRLGEP  
CDQLHVCDASQGLVCQPGAGPGGRGALCLLAEDDSSCEVNGRLYREGETFQPHCSIRCRCEDG  
GFTCVPLCSEDVRLPSWDCPHPRRVEVLGKCCPEWVCGQGGLGTQPLPAQGPQFSGLVSSLP  
PGVPCPEWSTAWGPCSTTCGLGMATRVSNQNRFCRLETQRRLLCLSRPCPPSRGRSPQNSAF

**Important features:**

Signal sequence:

Amino acids 1-23

**N-myristoylation sites:**

Amino acids 3-9;49-55;81-87;85-91;126-132;164-170;166-172;  
167-173;183-189;209-215

**Insulin-like growth factor binding proteins signature:**

Amino acids 49-65

**von Willebrand C1 domain:**

Amino acids 107-124

**Thrombospondin 1 Homology Block:**

Amino acids 201-216

**IGF binding protein site:**

Amino acids 49-58



## **FIGURE 322**

MMGLSLASAVLLASLLSLHLGTATRGSDISKTC CFQYSHKPLPWTWVRSYEFTSNSCSQRAVI  
FTTKRGKKVCTHPRKKWVQKYISLLKTPKQL

**Important features:**

**Signal peptide:**

amino acids 1-23

**N-myristoylation sites.**

amino acids 3-9, 26-32

**Amidation site.**

amino acids 68-72

**Small cytokines (intecrine/chemokine).**

amino acids 23-88

Downloaded from www.sciencedirect.com

**FIGURE 323**

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAG**ATG**CAGGTGAGCAAGAGGATGCTGGCGGGGGGCGTGAGGAGCA  
 TGCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCTGGGCTCAGTGCTGTCAGGCTCGGCCA  
 CGGGCTGCCCCGCCCCGCTGCGAGTGCTCCGCCCAGGACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAG  
 TCCCCGAGGGCATCCCCACCGAGACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACG  
 AGTTCCGCCAGCTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCGCCT  
 TCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGAGCAACCGCCTGAAGCTCATCCCGCTAGGCGTCTTCA  
 CTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGATCGTTATCCTACTGGACTACATGTTTCAGG  
 ACCTGTACAACCTCAAGTCACTGGAGGTTGGCGACAATGACCTCGTCTACATCTCTACCGCGCCTTCAGCGGCC  
 TCAACAGCCTGGAGCAGCTGACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGC  
 ACGGCCTCATCGTCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGCTGTACC  
 GACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAGTGCCTCTACGGCCTCAACC  
 TGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCCCTACCTGGCCGTCCGCCACCTAGTCTATCTCC  
 GCTTCCTCAACCTCTCCTACAACCCCATCAGCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGG  
 AGATCCAGCTGGTGGGCGGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCCTCAACTACCTGCGCGTGC  
 TCAATGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGGAGACACTCA  
 TCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGGCGCCGCTGGCGGCTCAACTTCA  
 ACCGGCAGCAGCCACGTGCGCCACGCCGAGTTTGTCCAGGGCAAG**AG**AGTTCAAGGACTTCCCTGATGTGCTAC  
 TGCCCAACTACTTCACCTGCCGCCGCGCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCC  
 ACACGGTGCAGTTTGTGTGCCGGGCGGATGGCGACCCGCCGCCGATCCTCTGGCTCTACCCCGAAAGCACC  
 TGGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCTACGCCCAGGTAC  
 AGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGACTCCATGCCCGCCCACCTGCATGTGC  
 GCAGCTACTCGCCGACTGGCCCCATCAGCCCAACAAGACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGAG  
 AGGCCAACAGCACCCGCGCCACTGTGCCTTTCCTTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCT  
 TCATCTCTTTCCTGGGCGTCGTCTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACACAA  
 AGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCGACGCGCCCCGCAAGT  
 TCAACATGAAGATGATA**TGA**AGGCCGGGGCGGGGGCAGGGACCCCCGGGCGGCCGGGCAGGGGAAGGGCCTGGT  
 CGCCACCTGCTCACTCTCCAGTCTTCCACCTCCTCCCTACCCTTCTACACACGTTCTCTTCTCCCTCCCGCC  
 TCCGTCCCCTGCTGCCCCCGCCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGG  
 GGACCCACCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCAATAAT  
 TCAATAAAAAAGTTACGAACCTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTTATGAAAACCTGAAATAA  
 TAAAAAGAGAAAAAACTAAAAA

**FIGURE 324**

MQVSKRMLAGGVRSMPSPLLACWQPILLVLGSLSGSATGCPPRCECSAQDRAVLCHRKCFVAVPEGIPTETRL  
 LDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFFNNLFNLRTLGLRSNRLKLIPLGVFTGLSNLTKQDI  
 SENKIVILLDYMFDLYNLKSLEVGDNLDVYISHRAFSGLSLEQLTLEKCNLTSIPTEALSHLHGLIVLRRLHL  
 NINAIRDYSFKRLYLKVLKLEISHWPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPIS  
 TIEGSMHELLRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTLEESVFHSGVGNLETILDSNPLACDC  
 RLLWVFRRRWRLNFNRRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQVFVDEGHTVQFVCRADG  
 DPPPAILWLSPRKHLVSAKSNGRLTVPDGTLEVRYAQVQDNGTYLCIAANAGGNDSMPAHLHVRSYSPDWPHQP  
 NKTFAFISNQPGEGEANSTRATVPFPFDIKTLIIATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPR  
 KSDAGISSADAPRKFNMKMI

**Important features:****Signal sequence:**

amino acids 1-41

**Transmembrane domain:**

amino acids 556-578

**N-glycosylation site.**

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496,  
 505-509, 526-530, 542-546

**Casein kinase II phosphorylation site.**

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

**Tyrosine kinase phosphorylation site.**

amino acids 590-598

**N-myristoylation site.**

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326,  
 344-350, 359-365, 493-499, 503-509, 605-611

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 32-43



**FIGURE 325**

CCACGCGTCCGCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGAG  
GCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGAGG  
AGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAGGAG  
GAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAACGGAG  
AGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGAGTAGG  
AAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGGAAAGAC  
ACAGAGAGACGGGAGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTGGCTGCTT  
TGGCATTGTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCTGGAGGGAC  
AGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGGCAGGGGTCC  
CTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCTCGAGCGCTGGTACTC  
TGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGGACTGGTGGAGC  
TACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGCCTGGTGAATGCA  
GCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGAGCTGAAGAGGGTT  
CTTTATGACCCCTTTCTGCCCCCATTAAGGCTCAGCACTGGAGGAGAGAAGCTCCGGGGAACC  
TTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCTGTGGTCAATGTGTCT  
GGAGGTCCCCCTCCTTTACAGCCACCGACTCAGTGAAGTGC GGCTGCTGTTTGGAGCTCGCGAC  
GGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTGAGGTGCAGCTCATTCAC  
TTCAACCAGGAACTCTACGGGAATTTACAGCGCTGCCTCCCGCGGCCCAATGGCCTGGCCATT  
CTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTCAGTCGCCTCCTTAACCGC  
GACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTCTTCAAGACCTGAGCCTGGAG  
CTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGCTCTCTCAGCACCCCGCCCTGC  
TCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAATATCACCTCCCTTCAGATGCAC  
TCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCTTCCAGAGCCTCAGCGGTAACAGC  
CGGCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGCAACAGGGACCCCGGCACCCCGAG  
AGGCGCTGCCGAGGCCCAACTACCGCCTGCATGTGGATGGTGTCCCCCATGGTCGCTGAGAC  
TCCCCTTGAGGATTGCACCCGCCCCTCCTAAGCCTCCCCACAAGGCGAGGGGAGTTACCCCT  
AAAACAAAGCTATTAAAGGGACAGAATACTTA

**FIGURE 326**

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLCA  
VGKRQSPVDVELKRVLYDPFLPPLRLSTGGKELRGTYNTGRHVSFLPAPRPVVNVSGGPPLY  
SHRLSELRLFLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSLFVN  
VASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSLTPPCSETVTW  
ILIDRALNITSLQMHSRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPERRCRGP  
NYRLHVDGVPHGR

**Important features:****Signal peptide:**

Amino acids 1-23

**Transmembrane domain:**

Amino acids 177-199

**N-glycosylation sites:**

Amino acids 118-122;170-174;260-264

**Eukaryotic-type carbonic anhydrases proteins:**

Amino acids 222-271;128-165;45-93

**FIGURE 327**

GGACTAATCTGTGGGAGCAGTTTATTCCAGTATCACCCAGGGTGCAGCCACACCAGGACTGTGTTGAAGGGTGT  
TTTTTCTTTTAAATGTAATACCTCCTCATCTTTCTTCTTACACAGTGTCTGAGAACATTTACATTATAGATAA  
GTAGTACATGGTGGATAACTTCTACTTTTAGGAGGACTACTCTTCTGACAGTCCTAGACTGGTCTTCTACACT  
AAGACACCATGAAGGAGTATGTGCTCCTATTATTCCCTGGCTTTGTGCTCTGCCAAACCCTTCTTTAGCCCTTCAC  
ACATCGCACTGAAGAATATGATGCTGAAGGATATGGAAGACACAGATGATGATGATGATGATGATGATGATG  
ATGATGATGAGGACAACTCTCTTTTCCAACAAGAGAGCCAAGAAGCCATTTTTTTCCATTTGATCTGTTTCCAA  
TGTGTCCATTTGGATGTCAGTGCTATTACAGAGTTGTACATTGCTCAGATTTAGGTTTGACCTCAGTCCCAACCA  
ACATTCATTTGATACTCGAATGCTTGATCTTCAAAACAATAAAATTAAGGAAATCAAAGAAATGATTTTAAAG  
GACTCACTTCACTTTATGGTCTGATCCTGAACAACAACAAGCTAACGAAGATTACCCAAAAGCCTTTCTAACCA  
CAAAGAAGTTGCGAAGGCTGTATCTGTCCCACAATCACTAAGTGAAATACCACTTAATCTTCCCAAATCATTAG  
CAGAACTCAGAATTCATGAAAAATAAGTTAAGAAAAATACAAAAAGGACACATTCAAAGGAATGAATGCTTTACACG  
TTTTGGAAATGAGTGCAAACCCTCTTGATAATAATGGGATAGAGCCAGGGGCATTTGAAGGGGTGACGGTGTTC  
ATATCAGAATTCGAGAAGCAAACTGACCTCAGTTCTTAAAGGCTTACCACCAACTTTATTGGAGCTTCACTTAG  
ATTATAATAAAATTTCAACAGTGGAACCTTGAGGATTTTAAACGATACAAAGAACTACAAAGGCTGGGCCTAGGAA  
ACAACAAAATCACAGATATCGAAAATGGGAGTCTTGCTAACATACCACGTGTGAGAGAAAATACATTTGAAAAACA  
ATAAACTAAAAAAAATCCCTTCAGGATTACCAGAGTTGAAATACCTCCAGATAATCTTCCTTCATTCTAATTCAA  
TTGCAAGAGTGGGAGTAAATGACTTCTGTCCAACAGTGCCAAAGATGAAGAAATCTTTATACAGTGCAATAAGTT  
TATTCAACAACCCGGTGAAATACTGGGAAATGCAACCTGCAACATTTTCGTTGTGTTTTGAGCAGAATGAGTGTC  
AGCTTGGGAACCTTTGGAATGTAATAATTAGTAATTGGTAATGTCCATTTAATATAAGATTCAAAAATCCCTACAT  
TTGGAATACTTGAACCTCTATTAATAATGGTAGTATTATATATACAAGCAAATATCTATTCTCAAGTGGTAAGTCC  
ACTGACTTATTTTATGACAAGAAATTTCAACGGAATTTTGCCAACTATTGATACATAAGGGGTGAGAGAAAACA  
AGCATCTATTGCAGTTTCCTTTTTGCGTACAAATGATCTTACATAAATCTCATGCTTGACCATTCCCTTCTTCAT  
AACAAAAAAGTAAGATATTTCGGTATTTAACACTTTGTTATCAAGCACATTTTAAAAAGAACTGTACTGTAAATGG  
AATGCTTGACTTAGCAAAATTTGTGCTCTTTCATTTGCTGTTAGAAAAACAGAATTAACAAAGACAGTAATGTGA  
AGAGTGCATTACACTATTCTTATTCTTTAGTAACTTGGGTAGTACTGTAATATTTTTAATCATCTTAAAGTATGA  
TTTGATATAATCTTATTGAAATTACCTTATCATGTCTTAGAGCCCGTCTTTATGTTTAAAACTAATTTCTTAAAA  
TAAAGCCTTCAGTAAATGTTCAATTACCAACTTGATAAATGCTACTCATAAGAGCTGGTTTGGGGCTATAGCATAT  
GCTTTTTTTTTTTTTAATTATTACCTGATTTAAAAATCTCTGTAAAAACGTGTAGTGTTTCATAAAATCTGTAAC  
CGCATTTTAAATGATCCGCTATTATAAGCTTTTAAATAGCATGAAAATTGTTAGGCTATATAACATTGCCACTTCAA  
CTCTAAGGAATATTTTTGAGATATCCCTTTGGAAGACCTTGCTTGGAAGAGCCTGGACACTAACAAATTCTACACC  
AAATTGTCTCTTCAAATACGTATGGACTGGATAACTCTGAGAAACACATCTAGTATAACTGAATAAGCAGAGCAT  
CAAATTAAACAGACAGAAACCGAAAGCTCTATATAAATGCTCAGAGTCTTTTATGTATTTCTTATTGGCATTCAA  
CATATGTAAATCAGAAAACAGGGAAATTTTCATTAAAAATATTGGTTTGAAAT

**FIGURE 328**

MKEYVLLLFLALCSAKPFFSPSHIALKNMMLKDMEDTDDDDDDDDDDDDDEDNSLFPTREPRS  
HFFPFDLFPMCPFGCQCYSRVVHCSDLGLTSVPTNIPFDTRMLDLQNNKIKEIKENDFKGLTS  
LYGLILNNNKLTKEHPKAFLTTKKLRRLYLSHNQLSEIPLNLPKSLAELRIHENKVKKIQKDT  
FKGMNALHVLEMSANPLDNNNGIEPGAEGVTVFHIRIAEAKLTSVPKGLPPTLLELHLDYNKI  
STVELEDFKRYKELQRLGLGNNKITDIENGLANIPRVREIHLENNKLLKKIPSGLPPELKYLQI  
IFLHSNSIARVGVNDFCPTVPKMKKSLYSAISLFNNPVKYWEMQPATFRCVLSRMSVQLGNFGM

**Important features:****Signal sequence.**

amino acids 1-15

**N-glycosylation site.**

amino acids 281-285

**N-myristoylation sites.**

amino acids 129-135, 210-216, 214-220, 237-243, 270-276, 282-288

**Leucine zipper pattern.**

amino acids 154-176

**FIGURE 329**

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGCA  
TCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGTGT  
TTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACC**ATG**ATCACTGGTGTGTT  
CAGCATGCGCTTGTGGACCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACCAGCG  
GCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGCCTGCT  
GAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCCGCTCCC  
GCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACCTCAGTTTGA  
TTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCTCAATACCA  
TGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCATGCAGCAAAT  
GTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTCTTTCACCAAC  
CTTCAACCCACAGGAGGTCTTTATTTCGTTCCACTAACATTTTTTCGGAATCTGGAGTCCACCCG  
TTGTTTGCTGGCTGGGCTTTTCCAGTGTCAGAAAGAAGGACCCATCATCATCCACACTGATGA  
AGCAGATTTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGAGGCAGAGAACCAG  
AGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTGAAAAGGTGAAGGA  
CAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCTGGACAACGTGGCTGC  
CGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTGCACGGATGATCGAACA  
GAGAGCTGTGGACACATCCTTGTACATACTGCCCCAAGGAAGACAGGGAAAGTCTTCAGATGGC  
AGTAGGCCCCATTCTCCACATCCTAGAGAGCAACCTGCTGAAAGCCATGGACTCTGCCACTGC  
CCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATGTGACCTTCATACCGCTCTT  
AATGACCCTGGGGATTTTTTGACCACAAATGGCCACCGTTTGCTGTTGACCTGACCATGGAAC  
TTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTATTACCACGGGAAGGAGCAGGT  
GCCGAGAGGTTGCCCTGATGGGCTCTGCCCCTGGACATGTTCTTGAATGCCATGTCAGTTTA  
TACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAAACCTCAGGTGATGGAAGTTGGAAA  
TGAAGAG**TAA**CTGATTTATAAAAGCAGGATGTGTTGATTTTAAATAAAGTGCCTTTATACAATG

**FIGURE 330**

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMQVVFRHGARSPLKPLPLEEQVE  
WNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQLTKVGMQQMFALGERLRKNYVEDIPFL  
SPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEGP I I IHTDEADSEVLYPNYQSCWSLRQTRGRRTASLQ  
PGISED LKKVKDRMGIDSSDKVDF FILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESIQMA  
VGPFLLHILES NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEWFVQ  
LYYHGKEQVPRGCPDGLCPLDMFLNAMS VYTLSP EKYHALCSQTQVMEVGNEE

**Important features:****Signal sequence:**

amino acids 1-23

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 218-222

**Casein kinase II phosphorylation site.**

amino acids 87-91, 104-108, 320-324

**Tyrosine kinase phosphorylation site.**

amino acids 280-288

**N-myristoylation site.**

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

**Amidation site.**

amino acids 216-220

**Leucine zipper pattern.**

amino acids 10-32

**Histidine acid phosphatases phosphohistidine signature.**

amino acids 50-65

**FIGURE 331**

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACATTTTCCCTTTCCT  
 AACAAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTTGCTGGAGAAGAAAGGGCTGAGGGCAG  
 AGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTGCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGAT  
 GCAGCGGAGTGAGGTGATGGAAGTCTAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGA  
 CCTGGAGAGCCTGGGGGAGGGCCTGCCTAACAAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGATAAG  
 ACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCCTAATATCAAATTGACTGGCTGGGTGAACTTCAACAGCCT  
 TTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAATAGAGATGCTTTGTAAATAAAATTTTAA  
 AAAAAGCAAGTATTTTATAGCATAAAGGCTAGAGACCAAATAGATAACAGGATTCCCTGAACATTCTAAGAGG  
 GAGAAAGTATGTTAAAAATAGAAAAACCAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACC  
 CTGGGTGAGGCCAGCCTCTTTGCTCCTCCCGAAATTATTTTTGGTCTGACCACTCTGCCTTGTGTTTTGCAGAA  
 TCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCTCCTCACCGCCGCCCTCTC  
 AGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGGACAGCCGCTCTGTGGTCCTGCTCTCAGTGGTC  
 TGGGTGCTGCTGGCCCCCCCCAGCAGCCGGCATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACC  
 TTCAACCCTTGACCGTCCACCAAGGGACGGGGGCCGTCTATGTGGGGCCATCAACCGGGTCTATAAGCTGACA  
 GGCAACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCGCCCTCATC  
 GTGCAGCCCTGCAGCGAAGTGCTCACCCCTACCAACAATGTCAACAAGCTGCTCATCATTGACTACTCTGAGAAC  
 CGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGTCTGCAAGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTG  
 GAGCCATCCCACAAGAAGGAGCACTACCTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGC  
 TCTGAGGGTGAGGATGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCCTGTCC  
 AGCCGGAAGCTGCCCCGAGACCCTGAGTCTCAGCCATGCTCGACTATGAGCTACACAGCGATTTTGTCTCCTCT  
 CTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACATCTTCTACATCTACGGCTTTGCTAGT  
 GGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCCGAGACCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTC  
 TTCTACACCTCACGCATCGTGCGGCTCTGCAAGGATGACCCCAAGTTCCACTCATACTGTCCCTGCCCTTCGGC  
 TGCACCCGGGCGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCCCAG  
 GCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAGCAGTATCACCACCCG  
 CCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTTGAGATCAAGGAGCGCCTGCAGTCTCTGC  
 TACCAGGGCGAGGGCAACCTGGAGCTCAACTGGCTGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCC  
 ATCGATGATAACTTCTGTGGACTGGACATCAACCAGCCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTG  
 TACACCACCAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTGTGGGG  
 ACTAAGAGTGGAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCCATTACCTCCTCAGCAAA  
 GAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAGGCAACTTTATTTTCTTGGGGAACAAAGGTGA  
 AATGGGGAGGTAAGAAGGGGTAAATTTTGTGACTTAGCTTCTAGCTACTTCCCTCCAGCCATCAGTCATTGGGTAT  
 GTAAGGAATGCAAGCGTATTTCAATATTTCCCAAACCTTAAGAAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

**FIGURE 332**

MGTLGQASLFAPPGNYFWSHDHSAFCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWPR  
ALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENDRDWTFNHLLTVHQGTGAVYVGAINRVYK  
LTGNLTIQVAHKTGPEEDNKSRYPLIVQPCSEVLTLTNNVNKLLIIDYSENRLACGSLYQG  
VCKLLRLDDLFIIVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDVGKQDYFPTL  
SSRKLPRDPRESSAMLDYELHSDVFSSLIKIPSDTLALVSHFDIFYIYGFASSGGFVYFLTVQPE  
TPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKPGDSLALQ  
AFNITSQDDVLFALFSGKQKQYHHPDDSALCAFPRAINLQIKERLQSCYQGEGLNLELNWLL  
GKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNGYSVVFVGT  
KSGKLKKVRVYEFRCSSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

**Important features:****Signal sequence:**

amino acids 1-32

**Transmembrane domain:**

amino acids 71-87

**N-glycosylation site.**

amino acids 130-134, 145-149, 217-221, 381-385

**Casein kinase II phosphorylation site.**

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,  
384-388, 471-475, 481-485, 530-534

**N-myristoylation site.**

amino acids 220-226, 319-325, 353-359, 460-466, 503-509



**FIGURE 333**

GCTGAGTCTGCTGCTCCTGCTGCTGCTGCTCCAGCCTGTAACTGTGCCTACACCACGCCAGG  
 CCCCCCAGAGCCCTCACCACGCTGGGCGCCCCCAGAGCCACACCATGCCGGGCACCTACGC  
 TCCCTCGACCACACTCAGTAGTCCCAGCACCCAGGGCCTGCAAGAGCAGGCACGGGCCCTGAT  
 GCGGGACTTCCCGCTCGTGGACGGCCACAACGACCTGCCCCCTGGTCCTAAGGCAGGTTTACCA  
 GAAAGGGCTACAGGATGTTAACCTGCGCAATTTTCAGCTACGGCCAGACCAGCCTGGACAGGCT  
 TAGAGATGGCCTCGTGGGCGCCCAGTTCTGGTCAGCCTATGTGCCATGCCAGACCCAGGACCG  
 GGATGCCCTGCGCCTCACCTGGAGCAGATTGACCTCATAACGCCGCATGTGTGCCTCCTATTC  
 TGAGCTGGAGCTTGTGACCTCGGCTAAAGCTCTGAACGACACTCAGAAATTGGCCTGCCTCAT  
 CGGTGTAGAGGGTGGCCACTCGCTGGACAATAGCCTCTCCATCTTACGTACCTTCTACATGCT  
 GGGAGTGCCTACCTGACGCTCACCCACACCTGCAACACACCCTGGGCAGAGAGCTCCGCTAA  
 GGGCGTCCACTCCTTCTACAACAACATCAGCGGGCTGACTGACTTTGGTGAGAAGGTGGTGGC  
 AGAAATGAACCGCCTGGGCATGATGGTAGACTTATCCCATGTCTCAGATGCTGTGGCACGGCG  
 GGCCCTGGAAGTGTACAGGCACCTGTGATCTTCTCCCACTCGGCTGCCCGGGGTGTGTGCAA  
 CAGTGCTCGGAATGTTCTGATGACATCCTGCAGCTTCTGAAGAAGAACGGTGGCGTCGTGAT  
 GGTGTCTTTGTCCATGGGAGTAATACAGTGCAACCCATCAGCCAATGTGTCCACTGTGGCAGA  
 TCACTTCGACCACATCAAGGCTGTGATTGGATCCAAGTTCATCGGGATTGGTGGAGATTATGA  
 TGGGGCCGGCAAATTCCCTCAGGGGCTGGAAGACGTGTCCACATACCCGGTCTGATAGAGGA  
 GTTGCTGAGTCGTGGCTGGAGTGAGGAAGAGCTTCAGGGTGTCTTTCGTGGAAACCTGCTGCG  
 GGTCTTCAGACAAGTGGAAAAGGTACAGGAAGAAAACAAATGGCAAAGCCCCTTGGAGGACAA  
 GTTCCCGGATGAGCAGCTGAGCAGTTCTTGCCACTCCGACCTCTCACGTCTGCGTCAGAGACA  
 GAGTCTGACTTCAGGCCAGGAACTCACTGAGATTCCCATACTGGACAGCCAAGTTACCAGC  
 CAAGTGGTCAGTCTCAGAGTCCTCCCCCACATGGCCCCAGTCCTTGCAGTTGTGGCCACCTT  
 CCCAGTCCTTATTCTGTGGCTCTGATGACCCAGTTAGTCCTGCCAGATGTCACTGTAGCAAGC  
 CACAGACACCCACAAAGTTCCCCTGTTGTGCAGGCACAAATATTTCTGAAATAAATGTTTT  
 GGACATAG

**FIGURE 334**

MPGTYAPSTTLSSPSTQGLQEQARALMRDFPLVDGHNDLPLVLRQVYQKGLQDVNLRNFSYGO  
TSLDRLRDGLVGAQFWSAYVPCQTQDRDALRLTLEQIDLIIRMCASYSELELVTSAKALNDTQ  
KLACLIGVEGGHSLDNSLSILRTFYMLGVRYLTLTHTCNTPWAESSAKGVHSFYNNISGLTDF  
GEKVVAEMNRLGMMVDLSHVSDAVARRALEVSQAPVIFSHSAARGVCNSARNVPDDILQLLKK  
NGGVVMVSLSMGVIQCNPSANVSTVADHFDHIKAVIGSKFIGIGGDYDGAGKFPQGLEDVSTY  
PVLIEELLSRGWSEELQGVLRGNLLRVFRQVEKVQEEKWQSPLEDKFPDEQLSSSCHSDLS  
RLRQRQSLTSGQELTEIPIHWTAKLPKWSVSESSPHMAPVLAVVATFPVLILWL

**Important features:****N-glycosylation sites.**

amino acids 58-62, 123-127, 182-186, 273-277

**N-myristoylation sites.**

amino acids 72-78, 133-139, 234-240, 264-270, 334-340, 389-395

**Renal dipeptidase active site.**

amino acids 134-157



**FIGURE 336**

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLEW  
KKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEEDTV  
TLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQSTNSS  
YTMNTKTGTLQFNTVSKLDTGEYSCEARN SVGYRRCPGKRMQVDDLNISGIIAAVVVVALVIS  
VCGLGVCYAQRKGYFSKETS FQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQEF

**Important features:****Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 130-144, 238-258

**N-glycosylation site.**

amino acids 98-102, 187-191, 236-240, 277-281

**Casein kinase II phosphorylation site.**

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

**N-myristoylation site.**

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

**Amidation site.**

amino acids 226-230

**FIGURE 337**

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCGG  
CACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGCCA  
**TG**ATTTCCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTTGTTCTTGGGGCTGAGTG  
CCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAGGCGG  
TGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGCACGGGGAGGTGTCTTCATCCC  
AGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATCAGGTGT  
TGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCCATGCCCT  
CCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCCTACAGCTGCT  
CCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCTTAGAACTCA  
ATGTACTGGTTCTTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCATGTGGGGGCAA  
ACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCAGTGGGATCGGC  
AGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTGGGTCTTTAAGCC  
TCACCAACCTTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCAATGAGGTGGGCA  
CTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGCAGTGGTTGCTGGAG  
CTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCTTGTACCACCGCC  
GGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCCATTGCTCCCCGGACCC  
TGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCTTTTCTCTGTACCTCCG  
CACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCATTGACCCCCACGCCCAGTC  
TCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCCACGACAGATGGGGCCCACCCTCAACCAA  
TATCCCCCATCCCTGGTGGGGTTTCTTCTCTGGCTTGAGCCGCATGGGTGCTGTGCCTGTGA  
TGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTA**TGA**TGACCCCACTCATTGGCTAAAG  
GATTGGGGTCTCTCCTTCTATAAGGGTCACCTCTAGCACAGAGGCCTGAGTCATGGGAAAG  
AGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTCTTTACTGTGGGAAAACCATCTCA  
GTAAGACCTAAGTGTCAGGAGACAGAAGGAGAAGAGGAAGTGGATCTGGAATTGGGAGGAGC  
CTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTGCTGAAATTAGCTACTACCAAGAGT  
GAGGGGCAGAGACTTCCAGTCACTGAGTCTCCAGGCCCCCTTGATCTGTACCCACCCCTAT  
CTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGTATTGATATAACCTGTCAGGCTGGCTTG  
GTTAGGTTTTTACTGGGGCAGAGGATAGGGAATCTCTTATTAATAACATGAAATATGTGTT  
GTTTTCATTTGCAAATTTAAATAAAGATACATAATGTTTGTATGAAAAA

**FIGURE 338**

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGCVSSS  
QPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPYSC  
SVNVQDKQGKSRGHSIKTLELNVLPAPPSCRLQGVPHVGANVTLSQSPRSKPAVQYQWDR  
QLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPAAVVAG  
AVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPPWKSSDTISKNGTLSSVTS  
ARALRPPHGPFRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSRMGAVPV  
MVPAQSQAGSLV

**Important features:****Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 245-267

**N-glycosylation site.**

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

**N-myristoylation site.**

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,  
262-268, 308-314, 363-369, 364-370

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 164-175

**FIGURE 339**

GCGAGAACCTTTGCACGCGCACAACTACGGGGACGATTTCTGATTGATTTTGGCGCTTTCGATCCACCCTCCT  
 CCTTCTC**ATG**GGACTTTGGGGACAAAGCGTCCCGACCGCCTCGAGCGCTCGAGCAGGGCGCTATCCAGGAGCCA  
 GGACAGCGTCGGGAACCAGACCATGGCTCCTGGACCCCCAAGATCCTTAAGTTCGTCTTCATCGTCGCGGTTT  
 TGCTGCCGGTCCGGTTGACTCTGCCACCATCCCCCGGCAGGACGAAGTTCCTCAGCAGACAGTGGCCCCACAGC  
 AACAGAGGCGCAGCCTCAAGGAGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAATATACTGGAGCCTGTAACC  
 CGTGACAGAGGGTGTGGATTACACCATTGCTTCCAACAATTTGCCTTCTTGCCTGCTATGTACAGTTTGTAAAT  
 CAGGTCAAACAAATAAAAGTTCTGTACCACGACCAGAGACACCGTGTGTGAGTGTGAAAAAGGAAGCTTCCAGG  
 ATAAAAACTCCCCTGAGATGTGCCGGACGTGTAGAACAGGGTGTCCCAGAGGGATGGTCAAGGTGAGTAATTGTA  
 CGCCCCGGAGTGACATCAAGTGCAAAAATGAATCAGCTGCCAGTTCCACTGGGAAAAACCCAGCAGCGGAGGAGA  
 CAGTGACCACCATCCTGGGGATGCTTGCCTCTCCCTATCACTACCTTATCATCATAGTGGTTTTAGTCATCATTTT  
 TAGCTGTGGTTGTGGTTGGCTTTTCATGTCCGAAGAAATTCATTTCTTACCTCAAAGGCATCTGCTCAGGTGGTG  
 GAGGAGGTCCCGAACGTGTGCACAGAGTCTTTTCCGGCGCGCTTCATGTCTTCCAGAGTTCTTGGGGCGGAGG  
 ACAATGCCCGCAACGAGACCCCTGAGTAACAGATACTTGACGCCACCCAGGTCTCTGAGCAGGAAATCCAAGGTC  
 AGGAGCTGGCAGAGCTAACAGGTGTGACTGTAGAGTCCCGAGAGGAGCCACAGCGTCTGCTGGAACAGGCAGAAG  
 CTGAAGGGTGTGAGAGGAGGAGGCTGCTGGTTCCAGTGAATGACGCTGACTCCGCTGACATCAGCACCTTGCTGG  
 ATGCCTCGGCAACACTGGAAGAAGGACATGCAAGGAAACAATTCAGGACCAACTGGTGGGCTCCGAAAAGCTCT  
 TTTATGAAGAAGATGAGGCAGGCTCTGCTACGTCCTGCCTG**TGA**AAGAATCTCTTCAGGAAACCAGAGCTTCCCT  
 CATTTACCTTTTCTCCTACAAAGGGAAGCAGCCTGGAAGAAACAGTCCAGTACTTGACCCATGCCCAACAACT  
 CTACTATCCAATATGGGGCAGCTTACCAATGGTCTAGAACTTTGTTAACGCACTTGGAGTAATTTTATGAAAT  
 ACTGCGTGTGATAAGCAAACGGGAGAAATTTATATCAGATTCTTGGCTGCATAGTTATACGATTGTGTATTAAGG  
 GTCGTTTTAGGCCACATGCGGTGGCTCATGCCTGTAATCCAGCACTTTGATAGGCTGAGGCAGGTGGATTGCTT  
 GAGCTCGGGAGTTTGAGACCAGCCTCATCAACACAGTGAAGTCCATCTCAATTTAAAAAGAAAAAAGTGGTTT  
 TAGGATGTCAATCTTTGCAGTTCTTCATCATGAGACAAGTCTTTTTTCTGCTTCTTATATTGCAAGCTCCATCT  
 CTACTGGTGTGTGCATTTAATGACATCTAACTACAGATGCCGCACAGCCACAATGCTTTGCCTTATAGTTTTTTA  
 ACTTTAGAACGGGATTATCTTGTATTACCTGTATTTTCAGTTTCGGATATTTTGGCTTAATGATGAGATTATC  
 AAGACGTAGCCCTATGCTAAGTCATGAGCATATGGACTTACGAGGGTTCGACTTAGAGTTTGGCTTTAAGATA  
 GGATTATTGGGGCTTACCCCCACCTTAATTAGAGAAACATTTATATTGCTTACTACTGTAGGCTGTACATCTCTT  
 TTCCGATTTTTGTATAATGATGTAAACATGAAAAACTTTAGGAAATGCATTTATTAGGCTGTTTACATGGGTTG  
 CCTGGATACAAATCAGCAGTCAAAAATGACTAAAAATATAACTAGTGACGGAGGGAGAAATCCTCCCTCTGTGGG  
 AGGCACTTACTGCATTCCAGTTCTCCCTCCTGCGCCCTGAGACTGGACCAGGGTTTTGATGGCTGGCAGCTTCTCA  
 AGGGGCAGCTTGTCTTACTTGTAAATTTAGAGGTATATAGCCATATTTATTTATAAATAAATATTTATTTATTT  
 ATTTATAAGTAGATGTTTACATATGCCAGGATTTTGAAGAGCCTGGTATCTTTGGGAAGCCATGTGTCTGGTTT  
 GTCGTGCTGGGACAGTCATGGGACTGCATCTTCCGACTTGTCCACAGCAGATGAGGACAGTGAGAATTAAGTTAG  
 ATCCGAGACTGCGAAGAGCTTCTCTTTCAAGCGCCATTACAGTTGAACGTTAGTGAATCTTGAGCCTCATTTGGG  
 CTCAGGGCAGAGCAGGTGTTTATCTGCCCCGGCATCTGCCATGGCATCAAGAGGGAAGAGTGGACGGTGTGGG  
 AATGGTGTGAAATGGTTGCCGACTCAGGCATGGATGGGCCCCCTCTCGCTTCTGGTGGTCTGTGAAGTGAAGTCCCT  
 GGGATGCCTTTTAGGGCAGAGATTCTGAGCTGCGTTTTAGGGTACAGATTCCCTGTTTGGAGAGCTTGGCCCCCT  
 CTGTAAGCATCTGACTCATCTCAGAGATATCAATTTCTAAACACTGTGACAACGGGATCTAAAATGGCTGACACA  
 TTTGTCTTGTGTACGTTCCATTATTTTATTTAAAAACCTCAGTAATCGTTTTAGCTTCTTTCCAGCAAACCTCT  
 TCTCCACAGTAGCCAGTCGTGGTAGGATAAATTACGGATATAGTCATTCTAGGGGTTTCACTCTTTTCCATCTC  
 AAGGCATTGTGTGTTTTGTTCCGGGACTGGTTTTGGCTGGGACAAAGTTAGAAGTGCCTGAAGTTCGCACATTGAG  
 ATTGTTGTGTCCATGGAGTTTTAGGAGGGGATGGCCTTTCCGGTCTTCGCACTTCCATCCTCTCCCACTTCCATC  
 TGGCGTCCCAACACTTGTCCCCTGCACTTCTGGATGACACAGGGTGTGCTGCCTCCTAGTCTTTGCCTTTGCTG  
 GGCCTTCTGTGCAGGAGACTTGGTCTCAAAGCTCAGAGAGAGCCAGTCCGGTCCCAGCTCCTTTGTCCCTTCCCTC  
 AGAGGCCTTCCCTGAAGATGCATCTAGACTACCAGCCTTATCAGTGTTTAAGCTTATTCCTTTAACATAAGCTTC  
 CTGACAACATGAAATTTGTTGGGGTTTTTTGGCGTTGGTTGATTTGTTTAGGTTTTGCTTTATACCCGGGCCAAAT  
 AGCACATAACACCTGGTTATATATGAAATACTCATATGTTTATGACCAAAATAAATATGAAACCTCATRTTAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 340**

MGLWGQSVPTASSARAGRYPGARTASGTRPWLLDPKILKFVVFIVAVLLPVRVDSATIPRQDEVPPQTVAPQQQR  
RSLKEEECPAGSHRSEYTGACNPCTEGVDYTIASNNLPSCLLCTVCKSGQTNKSSCTTTRDTCQCEKGSFQDKN  
SPEMCRTCRTGCPRGMVKVSNCTPRSDIKCKNESAASSTGKTPAAEETVTTILGMLASPHYHLIIIVLVIILAV  
VVVGFSCRKKFISYLGICSGGGGGPERVHRVLFRRRSCPSRVPGAEDNARNETLSNRYLQPTQVSEQEIQQQEL  
AELTGVTVESPEEPQRLLEQAEAEGCQRRRLLPVNDADSADISTLLDASATLEEGHAKETIQDQLVGSEKLFYE  
EDEAGSATSCSCL

**Important features:****Transmembrane domains:**

amino acids 35-52, 208-230

**N-glycosylation sites.**

amino acids 127-131, 182-186, 277-281

**Glycosaminoglycan attachment site.**

amino acids 245-249

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 260-264

**N-myristoylation sites.**

amino acids 21-27, 86-92, 102-108, 161-167, 242-248, 270-276, 297-303, 380-386

**ATP/GTP-binding site motif A (P-loop).**

amino acids 185-193

**TNFR/NGFR cysteine-rich region.**

amino acids 99-139



**FIGURE 341**

GCCTCTGAATTGTTGGGCAGTCTGGCAGTGGAGCTCTCCCCGGTCTGACAGCCACTCCAGAGG  
CC**ATG**CTTCGTTTCTTGCCAGATTTGGCTTTCAGCTTCCTGTTAATTCTGGCTTTGGGCCAGG  
CAGTCCAATTTCAAGAATATGTCTTTCTCCAATTTCTGGGCTTAGATAAGGCGCCTTCACCCC  
AGAAGTTCCAACCTGTGCCTTATATCTTGAAGAAAATTTTCCAGGATCGCGAGGCAGCAGCGA  
CCACTGGGGTCTCCCGAGACTTATGCTACGTAAAGGAGCTGGGCGTCCGCGGGAATGTACTTC  
GCTTTCTCCCAGACCAAGGTTTCTTTCTTTACCCAAAGAAAATTTCCCAAGCTTCCTCCTGCC  
TGCAGAAGCTCCTCTACTTTAACCTGTCTGCCATCAAAGAAAGGGAACAGTTGACATTGGCCC  
AGCTGGGCCTGGACTTGGGGCCCAATTCTTACTATAACCTGGGACCAGAGCTGGAAGTGGCTC  
TGTTCTGGTTCAGGAGCCTCATGTGTGGGGCCAGACCACCCCTAAGCCAGGTAAAATGTTTG  
TGTTGCGGTCAAGTCCCATGGCCACAAGGTGCTGTTCACTTCAACCTGCTGGATGTAGCTAAGG  
ATTGGAATGACAACCCCCGGAAAAATTTTCGGGTATTCTCGGAGATACTGGTCAAAGAAGATA  
GAGACTCAGGGGTGAATTTTCAGCCTGAAGACACCTGTGCCAGACTAAGATGCTCCCTTCATG  
CTTCCCTGCTGGTGGTGACTCTCAACCCTGATCAGTGCCACCCTTCTCGGAAAAGGAGAGCAG  
CCATCCCTGTCCCCAAGCTTTCTTGTAAGAACCTCTGCCACCGTCACCAGCTATTCATTA  
TCCGGGACCTGGGTGGCACAAGTGGATCATTGCCCCAAGGGGTTCATGGCAAATTACTGCC  
ATGGAGAGTGTCCCTTCTCACTGACCATCTCTCTCAACAGCTCCAATTATGCTTTCATGCAAG  
CCCTGATGCATGCCGTTGACCCAGAGATCCCCAGGCTGTGTGTATCCCCACCAAGCTGTCTC  
CCATTTCCATGCTCTACCAGGACAATAATGACAATGTCATTCTACGACATTATGAAGACATGG  
TAGTCGATGAATGTGGGTGTGGG**TAG**GATGTCAGAAATGGGAATAGAAGGAGTGTTCTTAGGG  
TAAATCTTTTAATAAAACTACCTATCTGGTTTATGACCACTTAGATCGAAATGTC

**FIGURE 342**

MLRFLPDLAFSFLLILALGQAVQFQEYVFLQFLGLDKAPSPQKFQPVPIYILKKIFQDREAAAT  
TGVSRDLCYVKELGVRGNVLRFLPDQGFFLYPKKISQASSCLQKLLYFNLSAIKEREQLTLAQ  
LGLDLGPNSYYNLGPELELALFLVQEPHVWGQTPKPGKMFVLRVWPWPQGAVHFNLLDVAKD  
WNDNPRKNFGLFLEILVKEDRDSGVNFQPEDTCARLRCSLHASLLVVTLNPDQCHPSRKRRAA  
IPVPKLSCKNLCHRHQLFINFRDLGWHKWIIAPKGFMANYPCHGECPPFSLTISLNSSNYAFMQA  
LMHAVDPEIPQAVCIPTKLSPISMPLYQDNNDNVILRHYEDMVVDECGCG

**Important features:****Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 112-116, 306-310

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 96-100

**N-myristoylation site.**

amino acids 77-83

**TGF-beta family proteins.**

amino acids 264-299, 327-341, 345-364



**FIGURE 344**

MELVRRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTIE  
 RPENKSIRIIIFSIVQLDPDGSCSENIKVFDGTSSNGPLLQVCSKNDYVPVFESSSSTLTFQ  
 IVTDSARIQRTVFVFYFFSPNISIPNCGGYLDTLEGSTSPNYPKPHPELAYCVWHIQVEKD  
 YKIKLNFKEIFLEIDKQCKFDLAIYDGPSTNSGLIGQVCGRVPTTFESSSNSLTVVLDYD  
 NSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRPKLSNV  
 VEFVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNSTVEIYYI  
 TEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPNLVVFLDT  
 CRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYLQCKVLICD  
 SSDHQSRCNQGCVSRSKRDISSYKWKTDIIIGPIRLKRRDRSASGNSGFQHETHAEETPNQPFN  
 SVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

**Important features:****Signal sequence:**

amino acids 1-24

**Transmembrane domain:**

amino acids 571-586

**N-glycosylation site.**

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,  
 394-398, 419-423

**Casein kinase II phosphorylation site.**

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,  
 408-412, 463-467, 520-524, 556-560

**Tyrosine kinase phosphorylation site.**

amino acids 172-180, 407-415, 407-416, 519-528

**N-myristoylation site.**

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18



**FIGURE 346**

MGSRCALALAVLSALLCQVWSSGVFELKLQEFVNKKGLLGNRNCCRGGAGPPPCACRTFFRVC  
 LKHYQASVSPEPPCTYGSATPVLGVDSFSLPDGGGADSAFSNPIRFPFGFTWPGTFSLII EA  
 LHTDSPDDLATENPERLISRLATQRHLTVGEEWSQDLHSSGRDLDKYSYRFVFCDEHYYGEGCS  
 VFCRPRDDAFGHFTCGERGEKVCNPGWKGPYCTEPICLPGCDEQHGFCDKPGECKCRVGVWQGR  
 YCDECIRYPGCLHGTCQQPWQCNCQEGWGGLFCNQDLNYCTHHKPKNGATCTNTGQGSYTCS  
 CRPGYTGATCELGIDECDPSPCKNGGSCTDLENSYSCTCPPGFGYKICELSAMTCADGPCFNG  
 GRCSDSPDGGYSCRCVPVGYSGFNCEKKIDYCSSSPCSNGAKCVDLGDAYLCRCQAGFSGRHCD  
 DNVDDCASSPCANGGTCRDGVNDFSCTCPPGYTGRNCSAPVSRCEHAPCHNGATCHERGHRYV  
 CECARGYGGPNCQFLLPELPPGPAVVDLTEKLEGQGGPFPWVAVCAGVILVLMLLLGCAAVVV  
 CVRLRLQKHRPPADPCRGETETMNNLANCQREKDISVSIIGATQIKNTNKKADFHGDHSADKN  
 GFKARYPAVDYNLVQDLKGDDTAVRDAHSKRDTKCQPQGSSGEEKGTPPTLRGGEASERKRPD  
 SGCSTSKDTKYQSVYVISEEKDECVIATEV

**Important features:****Signal sequence:**

Amino acids 1-21

**Transmembrane domain:**

Amino acids 546-566

**N-glycosylation site:**

Amino acids 477-481

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 660-664

**Tyrosine kinase phosphorylation sites:**

Amino acids 176-185;252-261

**N-myristoylation sites:**

Amino acids 2-8;37-43;40-46;98-104;99-105;262-268;281-287;  
 282-288;301-307;310-316;328-334;340-344;378-384;387-393;512-518;  
 676-682;683-689;695-701

**Aspartic acid and asparagine hydroxylation sites:**

Amino acids 343-355;420-432;458-470

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 552-563

**EGF-like domain cysteine pattern signature:**

Amino acids 243-255;274-286;314-326;352-364;391-403;429-441;  
 467-479;505-517

**FIGURE 347**

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAACA  
 TCGTAGTCCACCCCTCCCCATCCCCAGCCCCGGGGATTAGGCTCGCCAGCGCCCAGCCAG  
 GGAGCCGGCCGGGAAGCGCG**ATG**GGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGCTGT  
 TCGCCTGCTGCTGGGCGCCCCGGCGGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCTGGACAT  
 CTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCACGAGG  
 ACTCATCCCTGCAATGGTCTAACCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGAGAGCCC  
 TTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGCATCAGCA  
 ATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGTGCGAACTG  
 CCAAGTCCCTCGTCACTGTGCTAGGAATTCACAGAAGCCCATCATCACTGGTTATAAATCTT  
 CATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAGCCTGCAGCCC  
 GGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCATACAGGAAGATC  
 CCAATGGTAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCCGGGAGGATGATG  
 GGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGACAGATCCACCTCTC  
 AACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCCTCCCCATCCTCGTG  
 AGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCCCCCAGCAGTACCTAT  
 GGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGTGCCCTGATCTTCCCTT  
 TCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAGCAACATGGGCAGCTACA  
 AGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCTCCTCCTCCAGCACCTACC  
 ACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTGCTCATCATGCTCATCTTCC  
 TTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACATGAGGCAAAAGGCTCCGACG  
 ATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCGGGCAGTCAGGAGGGGACGACA  
 AGAAGGAATATTTATC**TAG**AGGCGCCTGCCCCACTTCCTGCGCCCCCAGGGGCCCTGTGGGG  
 ACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAACCGCAGGGCCGCCCTCCCGCTT  
 GCTCCCCAGCCCACCCACCCCTGTACAGAATGTCTGCTTTGGGTGCGGTTTTGTACTCGGT  
 TTGGAATGGGGAGGGAGGAGGGCGGGGGGAGGGGAGGGTTGCCCTCAGCCCTTTCGGTGGCTT  
 CTCTGCATTTGGGTATTATTTATTTTGTAAACAATCCCAAATCAAATCTGTCTCCAGGCTGGA  
 GAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACAAACAAAAACA

**FIGURE 348**

MGAPAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQWS  
 NPAQQTLYFGEKRALRDNRIQLVTSTPHELSSISISNVALADEGEYTCSIFTMPVRTAKSLVTV  
 LGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQUEDPNGKTFT  
 VSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQKLLL  
 HCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKAYYTLN  
 VNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLTAEAKGSDDAPDADT  
 AIINAEGGQSGGDDKKEYFI

**Important features:****Signal sequence:**

amino acids 1-20

**Transmembrane domain:**

amino acids 331-352

**N-glycosylation site.**

amino acids 25-29, 290-294

**Casein kinase II phosphorylation site.**

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

**N-myristoylation site.**

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,  
 306-310, 334-340, 360-364, 385-389, 386-390

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18





**FIGURE 350**

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIPE  
 GVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTISR  
 AALAQLLKLEELHLDNDSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELRVDE  
 NRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPDLPGT  
 HLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNNPWFCD  
 CSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMARELNMNLLSCPTTTPGLPLFTPAPSTAS  
 PTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGGERVTPPISERIQLSIHFVNDTSIQVSW  
 LSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPLDAFNRAV  
 EDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVLLSVFCWHMH  
 KKGRTYSQKWKNRGRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKGD FRLQPIYTP  
 NGGINYTDCHIPNNMRYCNSSVPDLEHCHT

**Important features:****Signal peptide:**

amino acids 1-42

**Transmembrane domain:**

amino acids 542-561

**N-glycosylation site.**

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

**Casein kinase II phosphorylation site.**

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

**Tyrosine kinase phosphorylation site.**

amino acids 319-328

**N-myristoylation site.**

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,  
 522-528, 545-551, 633-639

**Amidation site.**

amino acids 581-585

**Leucine zipper pattern.**

amino acids 164-186

**Phospholipase A2 aspartic acid active site.**

amino acids 39-50



**FIGURE 352**

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDCSRKRLARLPEPLPSW  
 VARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANITLLSLAGNRIVEILPEHLKEFQSLETL  
 DLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGYFDNLANTLLVLKLNRRNISAIPPKMFKLPQLQHLELNRR  
 KIKNVDGLTFQGLGALKSLKMQRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAIN  
 RISPDWAEFCQKLSELDLTFNHL SRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSSLKTLDLKNNEISWTI  
 EDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDSLDAIMSLQGNAFSOMKKLQQLHLNTSSLLCDC  
 QLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDFPKPQITVQPETQSAIKGSNLSFICSAASS  
 SDSPMTFAWKDNEILLHDAEMENYAHRAQGGVEYTTILRLREVEFASEGKYQCVISNHFSSYSVKAKLTVN  
 MLPSFTKTPMDLTIRAGAMARLECAAVGHPAPQIAWQKDGTDFFPAARERRMHVMPEDDVFFIVDVKIEDIGVYS  
 CTAQNSAGSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPPKNLWTKDDSPVVTERHFFAAGNQ  
 LLIIVDSVDSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWATVGVVIAVVCVVGTSLV  
 WVVIYHTRRRNEDCSITNTDETNPADIPSYSLSQGTIADRQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGT  
 CHIDNSSEADVEAATDLFLCPFLGSTGPMYLGKNVYGSDFETYHTGCSPPDPRTVLMDHYEPSYIKKKECYPCSH  
 PSEESCERSFSNISWPSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDIFSANPEPASVASSNSFMGTFGKALRRPHL  
 DAYSSFGQPSDCQPRAFYLKAHSSPDLDGSEEDGKERTDFQENHICTFKQTLNRYRTPNFQSYDLDT

**Important features:****Signal sequence:**

amino acids 1-27

**Transmembrane domain:**

amino acids 808-828

**N-glycosylation site.**amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,  
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020**Glycosaminoglycan attachment site.**

amino acids 886-890

**Casein kinase II phosphorylation site.**amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,  
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,  
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,  
1073-1077, 1079-1083, 1081-1085**Tyrosine kinase phosphorylation site.**

amino acids 667-675

**N-myristoylation site.**amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,  
513-519, 588-594, 672-678, 683-687, 774-780, 933-939**Leucine zipper pattern.**

amino acids 58-80, 65-87



**FIGURE 354**

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYHL  
 FLHGNSLTRLPNEFANFYNAVSLHMENGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQTF  
 LGLDDLEYLQADFNLLRDIDPGAQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRGNRL  
 KTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ GKDL  
 NETTEQDLCPKKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTKIPGNW  
 QIKIRPTAAIATGSSRNKPLANSRPCPGGCSCDHIPGSGMKMNCNNRNVSSLADLKPKLSNVQ  
 ELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNYLDTLSRE  
 KFAGLQNLLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSLSKLSLHNN  
 YFMYLPVAGVLDQLTSIIQIDLHGPNWECSCTIVPFKQWAERLGSEVLMSDLKCETPVNFFRK  
 DFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVLVPGLLLVFVT  
 SAFTVVGMLVFILNRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADGAHRVYDCGSHS  
 LSD

**Important features:****Signal sequence:**

amino acids 1-15

**Transmembrane domain:**

amino acids 618-638

**N-glycosylation site.**

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 122-126, 646-650

**Casein kinase II phosphorylation site.**amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,  
608-612, 657-661, 666-670, 693-697**N-myristoylation site.**amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,  
354-360, 465-471, 493-499, 598-604, 603-609**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 337-348



**FIGURE 356**

MRPAFALCLLWQALWPGPGGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALST  
 VRAGAE LRAVLALLRAGPGPGGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLES  
 TLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGAASN  
 LSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIAD EIGARWDKLSGDVLCPCPGRYLR  
 AGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPPATATSPVP  
 QRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITPSGSVIS  
 KFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILTM TVLGLVKLCFHES PSSQPRKESMGPP  
 GLESDPEPAALGSSSAHCTNNGVKVGDCLDRDRAEGALLAESPLGSSDA

**Important features:****Signal sequence:**

amino acids 1-16

**Transmembrane domain:**

amino acids 399-418

**N-glycosylation site.**

amino acids 189-193, 381-385

**Glycosaminoglycan attachment site.**

amino acids 289-293

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 98-102, 434-438

**Casein kinase II phosphorylation site.**

amino acids 275-279, 288-292, 342-346, 445-449

**N-myristoylation site.**amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,  
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,  
477-483**Aspartic acid and asparagine hydroxylation site.**

amino acids 262-274



CCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGAAGA  
AGACTAAAAATGTTGTTTCCAATGTGGACACTGAAGAGACAAATTCTTATCCTTTTTTAACATAATCCTAATTTCC  
AAACTCCTTGGGGCTAGATGGTTTCTAAAACCTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG  
ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCCCTC  
ACCATTAACCATACACGACATCTCCCCAGCGTCTTTTTCACAGATCGGACCATTCTGGTAGAGATCGATTTTCAGA  
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAACCAATCATGTGCATCAAGAGGCTGCAGATTAACACCCAGAAGC  
TTTAGTGGACTCACTTATTTAAAAATCCCTTTACCTGGATGGAACACCGACTACTAGAGATACCGCAGGGCCCTCCCG  
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACAACATCTTTTTCCATCAGAAAAGAGAATCTAACGAACCTGGCC  
AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTATCGAAATCCTTGTTATGTTTCATATTCAATAGAGAAA  
GATGCCTTCTTAAACTTGACAAAGTTAAAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCCTACTGTT  
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC  
CTCAACCAATACAAATCTTTGACCTAAGTGGAAATGCCCTCGTTGTTATAATGCCCATTTCTCTGTGCGCGG  
TGTAATAATAATCTCCCTACAGATCCCTGTAAATGCTTTTGATGCGTGACAGAATTAAGAGTTTTACGTCTCA  
CACAGTAACCTCTCTTCAGCATGTGCCCCCAAGATGGTTTTAAGAACATCAACAACTCCAGGAAGTGGATCTGTCC  
CAAACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG  
TCTTTCAATTTTGAACCTCAGGTCTATCGTGCATCTATGAATCTATCACAGCATTTTCTTCACTGAAAAGCCTG  
AAAATTTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA  
AATCTTGAAGTCTTGATCTTGGCACAATACTTTATAAAAAATTGCTAACCTCAGCATGTTTAAACAAATTTAAAAGA  
CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAT  
GCCAGAACTTCTGTAGAAAGTTATGAACCCCAGGTCTCTGGAACATTACATTTATTCAGATATGATAAGTATGCA  
AGGAGTTGCAGATTCAAAAACAAGAGGCTTCTTTTCATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC  
TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCCTCTGATTTTCAGCATCTTCTTTTCTCAAATGCCCTG  
AATCTGTCAAGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTCACACCTTTAGCAGAGCTGAGATATTTG  
GACTTCTCCAACAACCGCTTGATTTACTCTCATTCAACAGCATTTGAAGAGCTTCACAACTGGAAGTTCTGGAT  
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAACTTTACCAAGAACCTAAAGGTT  
CTGCAGAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA  
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTTATTCAG  
AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTTCTTGCCTTCTGGAGTTTTTGTGTT  
ATGCCCTCAAATCTAAAGAAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTAGTTGGAAGAACTCCAGTGT  
CTAAAGAACCTGGAACCTTTGGACCTCAGCCACAACCAACTGACCACGTGCCCTGAGAGATTATCCAACCTGTTCC  
AGAAGCCTCAAGAATCTGATTTCTTAAGAATAATCAAACAGGAGTCTGACGAAGTATTTCTACAAGATGCCCTC  
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACAGGATTTCCGAGAAAATGCTCTC  
ACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTTCTGTGCACCTGTGATGCTGTGTGGTTTTGTCTGGTGG  
GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC  
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA  
TCTGTATCTCTCTTTCTCATGTTGATGATGACGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT  
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGATGTTGCTGATGATGCTTTTATGTGTAT  
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAA  
CATTTTAAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCAGAGCATA  
CAGCTTAGCAAAAAGACAGTGTTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTAAGATAGCATTTTAC  
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG  
TCCAAGTCTCAGCTCCGGAAGAGGCTCTGTGGGAGTCTGTCTCTTGAGTGGCCAAACACCCGCAAGCTCAC  
CCATACCTCTGCGCAGTGTCTAAAGAAGCCCTGGCCACAGACAATCATGTGCCCTATAGTCAGGTGTTCAAGGAA  
ACGGTCTAGCCCTCTTTTGCAAAACACAACCTGCCTAGTTTTACCAAGGAGAGGCCCTGGC

**FIGURE 358**

MVFPMWTLKRQILILEFNIILISKLLGARWFFKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGGI  
PTNTTNLTTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFSGL  
TYLKSLYLDGNQLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYRNP  
YVSYSIEKDAFLNLTKLVLSLKDNNVTAVPTVLPSTLTLEYLYNNMIAKIQEDDFNNLNQLQ  
ILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWFKNINK  
LQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFEHQVYRASMNLSQAFSSSLKSLKILRIR  
GYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKISPSGDSSE  
VGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQTLDSLKNS  
IFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLHSTAFEELHK  
LEVLDISSNSHYFQSEGITHMLNFTKNLKVQLKMMNDNDISSSTSRTMESESRLTLEFRGNH  
LDVLWREGDNRYLQLFKNLLKLEELDISKNLSFLPSGVFDGMPPNLKNLSLAKNGLKSFSWK  
KLQCLKNLETLDLSHNQLTTPPERLSNCSRSLKNLILKNNQIRSLTKYFLQDAFQLRYLDLSS  
NKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWWVNHTVETIPYLATDVTCVGP  
HKGQSVISLDLYTCELDLTNLILFSLISVSLFLMVMMTASHLYFWDVWYIYHFCKAKIKGYQ  
RLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEERDWLPGQPVLENLSQ  
SIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVILIFLEKPFQKSKFLQLRKRLC  
GSSVLEWPTNPQAHFYFWQCLKNALATDNHVAYSQVFKETV

**Important features:**

**Signal sequence:**

amino acids 1-26

**Transmembrane domain:**

amino acids 840-860



**FIGURE 360**

MHGSCSFLMLLLPLLLLLLVATTGPVGALTDEEKRLMVELHNLRYAQVSPTASDMLHMRWDEEL  
 AAFAKAYARQCVWGHNKERGRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQMC  
 GHYTQVVWAKTERIGCGSHFCEKLQGVETNIELLVLCNYEPPGNVKGKRPYQEGTPCSQCPSG  
 YHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRAEASDSRKMGTTPSSLATGIPAFVLVTEVSGSL  
 ATKALPAVETQAPTSLATKDPSPMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKSTHVPI  
 PKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEELPPSSEVLASVFPAQD  
 KPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVVSGLNSGP  
 GHVWGPLLGLLLLPLVLGIF

**Important features:****Signal sequence:**

amino acids 1-22

**N-glycosylation site.**

amino acids 114-118, 403-407, 409-413

**Glycosaminoglycan attachment site.**

amino acids 439-443

**Casein kinase II phosphorylation site.**

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

**N-myristoylation site.**

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

**Amidation site.**

amino acids 82-86, 172-176

**Peroxidases proximal heme-ligand signature.**

amino acids 287-298

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.**

amino acids 127-138

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.**

amino acids 160-172



## **FIGURE 362**

MRPLLVLALLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPG  
EKGEGRPGPLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSKRSESRVPPPSDAPLPFD  
RVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWPKPA  
SLSGGAMVRLEPEDQVWVQVGVDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA

### **Important features:**

#### **Signal sequence.**

amino acids 1-15

#### **N-myristoylation sites.**

amino acids 11-17, 68-74, 216-222

#### **Cell attachment sequence.**

amino acids 77-80

Downloaded from www.sciencedirect.com



**FIGURE 364**

MMWRPSVLLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAPLSDAPHDDAHGNFQYDHEAFLGRE  
VAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWSLAELRAWIAHTQQRHIRDSVSAAWDTYD  
TDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATREELT  
AFLHPPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAAWVQTERQQFRDF  
RDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFVGSQAT  
NYGEDLTRHHDEL

**Important features:****Signal sequence:**

amino acids 1-20

**N-glycosylation site.**

amino acids 140-144

**Casein kinase II phosphorylation site.**

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,  
291-295, 298-302

**N-myristoylation site.**

amino acids 263-269, 311-317

**Endoplasmic reticulum targeting sequence.**

amino acids 325-330



**FIGURE 365**

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCTAGTGGCCTGATCGCG**ATG**GGGACAAAG  
 GCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGGCA  
 TTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCTGTG  
 AAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCAAGGA  
 GACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGGTGACC  
 TTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATACACTTGT  
 ATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGTGCTTGTG  
 CCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCCACCATTGGGAACCGGGCAGTGCTG  
 ACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGATGGGATAGTG  
 ATGCCTACGAATCCCAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCTGAATCCCACA  
 ACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACAGCTGTGAGGCA  
 CGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCTGTGGAGCGGAAT  
 GTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAATCTTGTTTTTTGGC  
 ATCTGGTTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGACTTCGAGTAAGAAG  
 GTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAGACCTCGTCATTCTGT  
 GT**TGA**GCCTGGTTCGGCTCACCGCCTATCATCTGCATTTGCCTTACTCAGGTGCTACCGGACT  
 CTGGCCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTCTACACCCACAGGGCCC  
 CCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCCATCCTCCTTCATGCCCTC  
 CCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACCTTGTTTTAAAGTGTTTATTCCCCATTTCT  
 TTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCCTTCTAAGTAGACAGCAAAAA  
 TGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGCTGGCAGGGATCTTTGAATAGG  
 TATCTTGAGCTTGTTCTGGGCTCTTTCCTTGTGTACTGACGACCAGGGCCAGCTGTTCTAGA  
 GCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGGTGATGACACTGGGGTCCTTCCAT  
 CTCTGGGGCCCCACTCTTCTGTCTTCCCATGGGAAGTGCCACTGGGATCCCTCTGCCCTGTC  
 CTCTGAATACAAGCTGACTGACATTGACTGTGTCTGTGGAAAATGGGAGCTCTTGTTGTGGA  
 GAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAGGATTTAAAACCGCTGCTCTAAAGAAA  
 AGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCTGTAATCCCAGAGGCTGAGGCAGGCGGAT  
 CACCTGAGGTTCGGGAGTTTCGGGATCAGCCTGACCAACATGGAGAAACCCTACTGGAAATACAA  
 AGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTCCCAGCTGCTCAGGAGCCTGGCAACAAGAG  
 CAAAACCTCCAGCTCAAAAAAAAAAAAAAAAAA

**FIGURE 366**

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSLAYSGFSSPRVEWK  
FDQGDTTTRLVCYNNKITASYEDRVTFLLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKVKL  
IVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNSSYV  
LNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVA AVLVTLLILLGI  
LVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

**Important features:****Signal sequence:**

amino acids 1-27

**Transmembrane domain:**

amino acids 238-255

**N-glycosylation site.**

amino acids 185-189

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 270-274

**Casein kinase II phosphorylation site.**

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,  
193-197, 203-207, 287-291

**N-myristoylation site.**

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

**FIGURE 367**

GGGGAGAGGAATTGACCAIGTAAAGGAGACTTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG  
 AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGAAC  
 GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAA  
 TAAACAGAGAGTTAGACCCCGCGGGGGTTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCCTCC  
 CACCCCCAAAAAAGGATGATTGGAATGAAGAACCGAGGATTACAAAGAAAAAGTATGTTTCATTTTTCTC  
 TATAAAGGAGAAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT  
 GGTGTGGTGGTGTTCCTTTCTTTTTGAATTTCCACAAGAGGAGAGGAAATTAATAATACATCTGCAAAAGAA  
 TTTTTCAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCACAGTTGGA  
 TTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTTTTTTTTAAAT  
 TTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTGTTCTTAACCACCTGGATTTCATCTGGATGTTGCT  
 GTGATCAGTCTGAAATAACAACGTGTTGAATTCCAGAAGGACCAACACAGGATAAATTATGAATCTGGAACAAAGT  
 GACCTTACATCCACAGCAGATAATGATAGGTCTAGGTTTAAACAGGGCCCTATTGACCCCTGCTTGTGGTGCT  
 GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA  
 CCAGTTTCAGCAAGGTGATTTGTGTTGCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT  
 GCTGAACCTCCATGAGAACAATAACCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCATTGGAATCCT  
 ACAGTTTGTAGTAGGAACCATATCAGAACCATTGAAATTTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA  
 ACTCTTTGACAATCGTCTTACTACCATCCCGAATTGAGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGT  
 GCGAAACAACCCCATTTGAAAGCATCCCTTCTTATGCTTTTAAACAGAATTCCTTCTTTGCGCCGATAGACTTAGG  
 GGAATTTGAAAAGACTTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACCTTGAGGTATTTGAACCTTGC  
 CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACCTAGATGAGCTGGATCTTCTGGGAATCA  
 TTTATCTGCCATCAGGCCTGGCTCTTCCAGGGTTTGATGCACCTTCAAAAACCTGTGGATGATACAGTCCAGAT  
 TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACAACAATACTAAC  
 ATTACTGCCTCATGACCTCTTCACTCCCTTGTCATCATCTGAGCGGATACATTTACATCACAAACCTTGAACATG  
 TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTTGTGCCCGGTG  
 TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT  
 GATTGTGGAGCCCTTCGAGACCTCAATGTCTACTGAAGGCATGGCAGCTGAGCTGAAATGTCCGGCCTCCACATC  
 CCTGACATCTGTATCTTGGATTCTACTCCAAATGGAACAGTCAATGACATGAGGCGGTACAAAGTCCGGATAGCTGT  
 GCTCAGTGATGGTACGTTAAATTTTACAAATGTAACCTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA  
 TTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCAGCAACCCTACTCCTTTCTCTTACTTTTC  
 AACCCTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCC  
 AGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGGCACAAGGTGACAGAGAAAAACCTT  
 CACCATCCAGTGACTGATATAACACAGTGGGATCCCGAATTTGATGAGGTCATGAAGACTACCAAAATCATCAT  
 TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTATGCTGGTCATTTTCTACAAGATGAGGAAGCAGACCA  
 TCGGCAAAACCATCACGCCCCAACAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC  
 CATGGAAGGCCACTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT  
 CAACCAACAACAACAGTTAAACACAATAAATTCATATACACAGTTCAGTGCATGAACCGTTATTGATCCGAATGAA  
 CTCTAAAGACAATGTACAAGAGACTCAAAATCTAAACACTTACAGAGTTACAAAAACAAACAATCAAAAAAAA  
 GACAGTTTATTAATAATGACACAATGACTGGGTCTAAATCTACTGTTTCAAAAAGTGTCTTTACAAAAAACAA  
 AAAAGAAAAGAAATTTATTTATTAATAAATCTATTGTGATCTAAAGCAGACAAAA

**FIGURE 368**

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVICVRKNLREVPDGIS  
 TNTRLNLNHNQIQIIVKNSFKHLRHLEILQLSRNHIRTIEIGAFNGLANLNTLELFDNRLTTIPNGAFVYLSKL  
 KELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLSYISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELD  
 LSGNHLSAIRPGSFQGLMHLQKLWMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLH  
 NPWNCNCDILWLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAELKC  
 RASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTQDTGMYTCMVNSVGNNTASATLNVTAATTP  
 FSYFSTVTVETMEPSQDEARTDNNVGPTPVVDWETTNTTSLTPQSTRSTEKFTTIPVTDINSGIPGIDEVMKT  
 TKIIIGCFVAITLMAAVMLVIFYKMRKQHRQNHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYS  
 YKSPFNHTTTVNTINSIHSSVHEPLLIRMNSKDNVQETQI

**Important features:****Signal sequence:**

amino acids 1-44

**Transmembrane domain:**

amino acids 523-543

**N-glycosylation site.**

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446,  
 488-492, 606-610

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

**Casein kinase II phosphorylation site.**

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

**N-myristoylation site.**

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397,  
 422-428, 433-439, 531-537



**FIGURE 370**

MEGEEAEQPAWFHQPWPRGASDSAPPAGTMAQSRVLLLLLLLLPPQLHLGPVLAVRAPGFGRSG  
 GHSLSPREENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLPEH  
 TNHLSLQNNQLEKIYPEELSRLHRLETNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNKLT  
 APRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFGSSNVEVLI  
 LSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDNETFWK  
 LSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSNQLREQG  
 IHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYFLEELNLS  
 YNRITSPQVHRDAFRKLRLRLSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALARGALAGMA  
 QLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYLQNNKISAVP  
 ANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKDRGRLGKEKEE  
 EEEEEEEEEETR

**Important features:****Signal sequence:**

amino acids 1-48

**N-glycosylation site.**

amino acids 243-247, 310-314, 328-332, 439-443

**Casein kinase II phosphorylation site.**

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

**N-myristoylation site.**

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,  
 477-483, 498-502, 539-545, 548-554

**Leucine zipper pattern.**

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,  
 535-557

**FIGURE 371**

CACTTTCTCCCTCTCTTCTTTACTTTTCGAGAAACCGCGCTTCCGCTTCTGGTCGAGAGACCTCGGAGACCGCG  
 CCGGGGAGACGGAGGTGCTGTGGGTGGGGGGGACCTGTGGCTGCTCGTACCGCCCCCACCCTCCTCTTCTGCAC  
 TGCCGTCTCCGGAAGACCTTTTCCCTGCTCTGTTTCTTACCGAGTCTGTGCATCGCCCCGGACCTGGCCGG  
 GAGGAGGCTTGGCCGGCGGGAGATGCTCTAGGGGCGGCGGGAGGAGCGGCCGGCGGGACGGAGGGCCCCGGCAG  
 GAAG**ATG**GGCTCCCGTGGACAGGGACTCTTGCTGGCGTACTGCCTGCTCCTTGCCTTTGCCCTCTGGCCTGGTCCT  
 GAGTCGTGTGCCCCATGTCCAGGGGGAACAGCAGGAGTGGGAGGGGACTGAGGAGCTGCCGTGCGCTCCGGACCA  
 TGCCGAGAGGGCTGAAGAACAACATGAAAAATACAGGCCAGTCAAGACCAGGGGCTCCCTGCTTCCCGGTGCTT  
 GCGCTGCTGTGACCCCGGTACCTCCATGTACCCGGCGACCGCGTGCCCCAGATCAACATCACTATCTTGAAAGG  
 GGAGAAGGGTGACCGCGGAGATCGAGGCCCTCCAAGGGAAATATGGCAAAACAGGCTCAGCAGGGGGCCAGGGGCCA  
 CACTGGACCCAAAGGGCAGAAGGGCTCCATGGGGGGCCCTGGGGAGCGGTGCAAGAGCCACTACGCCGCCTTTTC  
 GGTGGGCGGGAAGAAGCCCATGCACAGCAACCACTACTACCAGACGGTGATCTTCGACACGGAGTTCTGTGAACCT  
 CTACGACCACCTTCAACATGTTTACCGGCAAGTTCTACTGCTACGTGCCCGGCCCTTACTTCTTCAGCCTCAACGT  
 GCACACCTGGAACCAAGGAGACCTACCTGCACATCATGAAGAACGAGGAGGAGGTGGTGATCTTGTTCGCGCA  
 GGTGGGCGACCGCAGCATCATGCAAAGCCAGAGCCTGATGCTGGAGCTGCGAGAGCAGGACCAGGTGTGGGTACG  
 CCTCTACAAGGGCGAACGTGAGAACGCCATCTTCAGCGAGGAGCTGGACACCTACATCACCTTCAGTGGCTACCT  
 GGTCAAGCACGCCACCGAGCCCT**TAG**CTGGCCGGCCACCTCCTTTCTCTCGCCACCTTCCACCCCTGCGCTGTGC  
 TGACCCACCGCCTCTTCCCCGATCCCTGGACTCCGACTCCCTGGCTTTGGCATTCACTGAGACGCCCTGCACAC  
 ACAGAAAGCCAAAGCGATCGGTGCTCCAGATCCCGCAGCCTCTGGAGAGAGCTGACGGCAGATGAAATCACCAG  
 GGCGGGGCACCCGCGAGAACCCTCTGGGACCTTCCGCGGCCCTCTCTGCACACATCCTCAAGTGACCCCGCACGG  
 CGAGACGCGGGTGGCGGCAGGGCGTCCCAGGGTGGCGCACCGCGGCTCCAGTCTTGAAATAATTAGGCAAAT  
 CTAAAGGTCTCAAAAGGAGCAAAGTAAACCGTGGAGGACAAAGAAAAGGGTTGTTATTTTTGTCTTTCCAGCCAG  
 CCTGCTGGCTCCCAAGAGAGAGGCCCTTTTCAGTTGAGACTCTGCTTAAGAGAAGATCCAAAGTTAAAGCTCTGGG  
 GTCAGGGGAGGGGCCGGGGCAGGAACTACCTCTGGCTTAATTCTTTTAAGCCACGTAGGAACCTTTCTTGAGGG  
 ATAGGTGGACCCTGACATCCCTGTGGCCTTGCCCAAGGGCTCTGCTGGTCTTTCTGAGTCACAGCTGCGAGGTGA  
 TGGGGGCTGGGGCCCCAGGCGTCAGCCTCCAGAGGGACAGCTGAGCCCCCTGCCTTGGCTCCAGGTTGGTAGAA  
 GCAGCCGAAGGGCTCCTGACAGTGGCCAGGGACCCCTGGGTCCCCAGGCCCTGCAGATGTTTCTATGAGGGGCAG  
 AGCTCCTTGGTACATCCATGTGTGGCTCTGCTCCACCCCTGTGCCACCCAGAGCCCTGGGGGGTGGTCTCCATG  
 CCTGCCACCCTGGCATCGGCTTTCTGTGCCGCTCCACACAAATCAGCCCCAGAAGGCCCCGGGGCCTTGGCTT  
 CTGTTTTTTTATAAAACACCTCAAGCAGCACTGCAGTCTCCCATCTCCTCGTGGGCTAAGCATCACCGCTTCCACG  
 TGTGTTGTGTTGGTTGGCAGCAAGGCTGATCCAGACCCCTTCTGCCCCACTGCCCTCATCCAGGCCCTTGACCA  
 GTAGCCTGAGAGGGGCTTTTTCTAGGCTTCAGAGCAGGGGAGAGCTGGAAGGGGCTAGAAAGCTCCCGCTTGTCT  
 GTTCTCAGGCTCCTGTGAGCCTCAGTCCCTGAGACCAGAGTCAAGAGGAAGTACAGTCCCAATCACCGTGTCA  
 GGATTCACTCTCAGGAGCTGGGTGGCAGGAGAGGCAATAGCCCCCTGTGGCAATTGCAGGACCAGCTGGAGCAGGG  
 TTGCGGTGTCTCCACGGTGTCTCTGCCCTGCCATGGCCACCCAGACTCTGATCTCCAGGAACCCCATAGCCCC  
 TCTCCACCTCACCCCATGTTGATGCCAGGGTCACTCTTGCTACCCGCTGGGCCCCCAAACCCCGCTGCCTCTC  
 TTCCTTCCCCCATCCCCACCTGGTTTTGACTAATCCTGCTTCCCTCTCTGGGCTGGCTGCCGGGATCTGGGG  
 TCCCTAAGTCCCTCTCTTTAAAGAACTTCTGCGGGTCAGACTCTGAAGCCGAGTTGCTGTGGGCGTGCCCGGAAG  
 CAGAGCGCCACACTCGCTGCTTAAGTCCCCAGCTCTTCCAGAAAACATTAAACTCAGAATTGTGTTTTCAA

**FIGURE 372**

MGSRGQGGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEQHEKYRPS  
QDQGLPASRCLRCCDPGTSMPATAVPQINITILKGEKGDGRGLQGKYGKTGSAGARGHTG  
PKGQKGSMGAPGERCKSHYAASFVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMTGKFYCYV  
PGLYFFSLNVHTWNQKETYLHIMKNEEEVVILFAQVGDRSIMQSQSLMLELREQDQVWVRLYK  
GERENAIIFSEELDTYITFSGYLVKHATEP

**Important features:****Signal sequence.**

amino acids 1-25

**N-glycosylation site.**

amino acids 93-97

**N-myristoylation sites.**

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

**Amidation site.**

amino acids 150-154

**Cell attachment sequence.**

amino acids 104-107



FIGURE 373

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCTCTGTCCCCAAGCCG  
TTCTAGACGCGGGAAAAATGCCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGGAA  
GCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAATGC  
ACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTTCAGAGG  
ATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAACCCAAAG  
ATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAGAGTTCT  
TCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATGTGGTTAA  
TGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACCTGGTTCTTCC  
TTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTAAAAAAGGATC  
CATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATATGTGGGTATGG  
AAGGAGGAATTGTCTTAAGTGAGAATCAATGAAAAGACTTAACAGCCTTCTCAATATCCCAG  
AAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAACAGCTAGCAGTTT  
GCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGAAAAGATGTATTTA  
ATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCCCAACCAGGTAGTAG  
AAGGCTGTTGTTTCAGATATGGCTGTTACTTTTAATGGACTGACTCCAAATCAGATGCATGTGA  
TGATGTATGGGGTATACCGCCTTAGGGCATTGCGGCATATTTTCAATGATGCATTGGTTTTCT  
TACCTCCAAATGGTTCTGACAATGACTTGAGAAAGTGGTAGAAAAGCGTGAATATGATCTTTGTA  
TAGGACGTGTGTTGTCAATTATTTGTAGTAGTAACTACATATCCAATACAGCTGTATGTTTCTT  
TTTCTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAGTCAGTAGTACATTTTTTAA  
TGAGGGTGGTTTTTTTTCTTTAAACACATGAACATTGTAAATGTGTTGGAAAGAAGTGTTTTA  
AGAATAATAATTTTGCAAATAAACTATTAATAAATATTATATGTGATAAATTCTAAATTATGA  
ACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTTAAAAAATTTTAACAGGTCTTTA  
GCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATTTGTGATTAAAGTAAAACCTTTTAG  
CTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCTAAGCCTCCCCAAGTTCCAATGGAT  
TTGCCTTCTCAAAATGTACAACCTAAGCAACTAAAGAAAATTAAAGTGAAAGTTGAAAAAT

**FIGURE 374**

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERMELSKSFRVYCIILV  
KPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRKAYKYAFDKYRDQYNWFFLARPTTFAI  
IENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGGIVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAV  
CLKYAGVFAENAEDADGKDVFNNTKSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFG  
HIFNDALVFLPPNGSDND

**Important features:****Signal sequence:**

amino acids 1-33

**N-glycosylation site.**

amino acids 121-125, 342-346

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 319-323, 464-468

**Casein kinase II phosphorylation site.**

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403,  
409-413, 473-477, 729-733, 748-752

**Tyrosine kinase phosphorylation site.**

amino acids 736-743

**N-myristoylation site.**

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564,  
651-657, 657-663, 672-672

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 14-25

**Cell attachment sequence.**

amino acids 247-250

FIGURE 375

GTTGTGTCCTTCAGCAAAACAGTGGATTAAATCTCCTTGCACAAGCTTGAGAGCAACACAAT  
 CTATCAGGAAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAAGAAGAAAAGAAGAAA  
 AAAAAATC**ATG**AAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCACGGG  
 GCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCCAA  
 AGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATTGACAA  
 CCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGACAAGTG  
 GTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCGAGATCCA  
 GAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAACCACCCAAA  
 GACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATTTCTTCAGATAT  
 CTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGACCAGAGCCTAC  
 GGTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGACGAATACTTGGA  
 AATTCAGGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGACGTGGC  
 CGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACCTATCCACCATACATTTCAGAAGCCAA  
 GGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCCTCAGCAGTCCCCTC  
 AGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAAGAAAGGGGTGAAAGT  
 GGAAAACAGACCTTTCTCTCAAACTCATCTTCTTCAATGTCTCTGAACATGACTATGGGAA  
 CTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGCATCATGCTATTTGGTCC  
 AGGCGCCGTCAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGGCTGCGTCTGGCTGCTGCC  
 TCTTCTGGTCTTGCACCTGCTTCTCAAATTT**TGA**TGTGAGTGCCACTTCCCCACCCGGGAAAG  
 GCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACACCGACAGCAACCAATCAGATA  
 TATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGAAATTTGAGGGAGGGGAACAAA  
 GAATACTTTGGGGGGAAAAGAGTTTTTAAAAAAGAAATTGAAAATTGCCTTGCAGATATTTAGG  
 TACAATGGAGTTTTCTTTCCCAAACGGGAAGAACACAGCACACCCGGCTTGGACCCACTGCA  
 AGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAAGGGCTCAGCCTCTCTGCCCACAGA  
 GTGCCCCCACGTGGAACATTCTGGAGCTGGCCATCCCAAATTCAATCAGTCCATAGAGACGAA  
 CAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGGGCACCTTTGGTAGACTGTGCCACCACG  
 CGGTGTGTTGTGAAACGTGAAATAAAAAAGAGCAAAAAAAAAA

376/550

## **FIGURE 376**

MKTIQPKMHNSISWAI FTGLAALCLFQGV PVRSGDATFPKAMDNVTVRQGESATLRCTIDNRV  
TRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPKTS  
RVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQ  
GITREQSGDYEC SASNDVAAPVVR RVKVTVNYPPYISEAKGTGVPVGQKGT LQCEASAVPSAE  
FQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGN YTCVASNKLGH TNASIMLFGPGA  
VSEVSNGTSRRAGCVWLLPLLVLHLLLKF

**Important features:**

**Signal peptide:**

amino acids 1-28

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 377**

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCCTTTAGATTGTGAA  
**ATGT**GGCTCAAGGTCTTCACAACTTTCCCTTTCCCTTTGCAACAGGTGCTTGCTCGGGGCTGAAG  
 GTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTGAGGCCCTCTACCTACCCGTCCAC  
 TATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCACACA  
 ATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACCAACAC  
 AAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCTGATGAA  
 GGCAATTACATCGTGAAGGTCAACATTGAGGGAAATGGAACCTCTATCTGCCAGTCAGAAGATA  
 CAAGTCACGGTTGATGATCCTGTCAAAAGCCAGTGGTGCAGATTCATCCTCCCTCTGGGGCT  
 GTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGGCTAGCTTAC  
 CAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTCTCCCCAAAC  
 AATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCTGCCTGGTGAGG  
 AACCCTGTGAGTGAAATGGAAAGTGATATCATTATGCCCATCATATATTATGGACCTTATGGA  
 CTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGTTGACCTTGGAGAG  
 GCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACTCCTGGATTAGGAGG  
 ACTGACAATACTACATATATCATTAAGCATGGGCCCTCGCTTAGAAGTTGCATCTGAGAAAGTA  
 GCCCAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAACCGGCAGGCAAGATGAA  
 ACTCATTTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTGACAGAAAGGAAAATCA  
 TTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTTGATTATATCCATGTGTCTTCTC  
 TTCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAACTAGAAGGCAGGCCAGAA  
 ACAGAATACAGGAAAAGCTCAAACATTTTCAGGCCATGAAGATGCTCTGGATGACTTCGGAATA  
 TATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTCAGGATTCCAAGCAGGTCTGTTCCA  
 GCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGTGTATGAAGTTATTCAGCACATC  
 CCTGCCCAGCAGCAAGACCATCCAGAG**TGA**ACTTTTCATGGGCTAAACAGTACATTCGAGTGAA  
 ATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAGTATATTAATCTGGAATCAGTGAAGA  
 AACCAGGACCAACACCTCTTACTCATTATTCCCTTTACATGCAGAATAGAGGCATTTATGCAAA  
 TTGAACTGCAGGTTTTTTCAGCATATACACAATGTCTTGCAACAGAAAAACATGTTGGGGAA  
 ATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGGGGAGAACGAAAGTGACAGGGGTTTCCT  
 CATAAGTTTTGTATGAAATATCTCTACAAACCTCAATTAGTTCTACTCTACACTTTCATATC  
 ATCAACACTGAGACTATCCTGTCTCACCTACAAATGTGGAACTTTACATTGTTTCGATTTTTTC  
 AGCAGACTTTGTTTTATTAAATTTTTTATTAGTGTTAAGAATGCTAAATTTATGTTTCAATTTT  
 ATTTCCAAATTTCTATCTTGTTATTTGTACAACAAAGTAATAAGGATGGTTGTCACAAAAACA  
 AAATATGCCTTCTCTTTTTTTTCAATCACCAGTAGTATTTTTTGAGAAGACTTGTGAACACTT  
 AAGGAAATGACTATTAAAGTCTTATTTTTTATTTTTTTCAAGGAAAGATGGATTCAAATAAATT  
 ATTCTGTTTTTGCTTTTAAAAAAAAAAAAAA

**FIGURE 378**

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVGRQALYLPVHYGFHTPASDIQIIWLFERPHTMPKYLLGSVNKS  
VVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQNGTSLASQKIQVTVDVPVKPVVQIHPPSGAVEY  
VGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFSPQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPII  
YYGPYGLQVNSDKGLKVGEVFTVDLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMD  
YVCCAYNNITGRQDETHFTVIIITSVGLEKLAQKGKSLSPASITGISLFLIISMCLLFLWKKYQPYKVIKQKLEG  
RPETERYKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRVSPASDCVSGQDLHSTVYEVIQHIIPAQQQDHPE

**Important features:****Signal sequence:**

amino acids 1-18

**Transmembrane domain:**

amino acids 341-359

**N-glycosylation site.**

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

**Casein kinase II phosphorylation site.**

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

**Tyrosine kinase phosphorylation site.**

amino acids 272-280

**N-myristoylation site.**

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

379/550

## **FIGURE 379**

ATAGTAGAAGAATGTCTCTGAAATTACTGGATGAGTTTCAGTCATACTTTCACATGGGCACAA  
TTTCACATTCAAGCTCCTTATCCTAGGCTAATTTTATATTATGTTAAATCACTTGTTTTTGTT  
CTCACGGCTTCCTGCCTGCTATAGGCATAATTACGAGGAAGCAGAAGTTCTCCAGAAGCAAGC  
GCACATGCGTTCCAAAATAAGAGCAAATTCGCTCTAAACACAGGAAAAGACCTGAAGCTTTAA  
TTAAGGGGTTACATCCAACCCCAGAGCGCTTTTGTGGGCACTGATTGCTCCAGCTTCTGCGTC  
ACTGCGCGAGGGAAGAGGGAAGAGGATCCAGGCGTTAGAC**CATG**TATAGACACAAAAACAGCTG  
GAGATTGGGCTTAAAATACCCACCAAGCTCCAAAGAAGAGACCCAAGTCCCCAAAACATTGAT  
TTCAGGGCTGCCAGGAAGGAAGAGCAGCAGCAGGGTGGGAGAGAAGCTCCAGTCAGCCCACAA  
GATGCCATTGTCCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCACCGCTGCCCT  
GCCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGGAATAAG  
GAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGACCTCCCAGGCCAGTGCCGG  
GCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGCACCCCCCAGCA  
ATCCGCGCGCCGGGACAGAATGCCCTGCAGGAAGTTCTTCTGGAAGACCTTCTCCTCCTGCAA  
**ATAG**

**FIGURE 380**

MYRHKNSWRLGLKYPPSSKEETQVPKTLISGLPGRKSSSRVGEKLQSAHKMPLSPGLLLLLLS  
GATATAALPLEGGPTGRDSEHMQEAAAGIRKSSLLTFLAWWFEWTSQASAGPLIGEEAREVARR  
QEGAPPQQSARRDRMPCRNFFWKTFSSCK

### Important features:

Transmembrane domain:

amino acids 51-69

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 35-39, 92-96

N-myristoylation sites.

amino acids 64-70, 75-81, 90-96

Amidation site.

amino acids 33-37



GGCGCCGGTGACACCGGGCGGGCTGAGCGCCTCCTGCGGCCGGCCCTGCGCGCCCCGGGCCCGCCG  
 GCGCCGCCACGCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCGGGCCCGCGCCCCGCGC  
 CCGCGCCAGGTGAGCGCTCCGCCCGCCGCGAGGCCCGCCCCGGGCCCGCCCCCGCCCCGCC  
 CGGCCGGCGGGGAACCGGGCGGATTCTCTGCGCGTCAAACCACCTGATCCCATAAAACATTC  
 ATCCTCCC GGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCCGCCGCTCGCCCTG  
 TGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGCGGAGCG  
 CGCCGAGCCTCGTCCC GCGGCCGGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGATGCGGAC  
 CCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTCAGTCCCCGACGCGCCCCGCCCA  
 ACCCCTACG**ATGA**AGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTGCTGTGGCTG  
 CAGGCCTGGCAGGTGGCAGCCCCATGCCCAGGTGCCTGCGTATGCTACAATGAGCCCAAGGTG  
 ACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTGCTGCCAGCCAG  
 CGCATCTTCCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTCCGTGCCTGCCGC  
 AACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGCGGCTGCCTTCACT  
 GGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCCGGTCTGTGGACCCT  
 GCCACATTCCACGGCCTGGGCGCCTACACACGCTGCACCTGGACCGCTGCGGCCTGCAGGAG  
 CTGGGCCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTACCTGCAGGACAACGCG  
 CTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCACACACCTCTTCCTGCAC  
 GGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTGCACAGCCTCGACCGTCTC  
 TACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTTCCGTGACCTTGGCCGCCTC  
 ATGACACTCTATCTGTTTTGCCAACAACTCTATCAGCGCTGCCCACTGAGGCCCTGGCCCCCTG  
 CGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTGTGTGACTGCCGGGCACGCCCA  
 CTCTGGGCCTGGCTGCAGAAAGTTCCGCGGCTCCTCCTCCGAGGTGCCCTGCAGCCTCCCGCAA  
 CGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATGACCTGCAGGGCTGCGCTGTGGCC  
 ACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACCGATGAGGAGCCGCTGGGGCTTCCC  
 AAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGTACTGGAGCCTGGAAGACCAGCTTCG  
 GCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTGACAGCCCGCCGGGCAACGGCTCTGGC  
 CCACGGCACATCAATGACTCACCCCTTTGGGACTCTGCCTGGCTCTGCTGAGCCCCCGCTCACT  
 GCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTTCCCCACCTCGGGCCCTCGCCGGAGGCCA  
 GGCTGTTACGCAAGAACCGCACCCGCAGCCACTGCCGTCTGGGCCAGGCAGGCAGCGGGGT  
 GCGGGGACTGGTGACTCAGAAGGCTCAGGTGCCCTACCCAGCCTCACCTGCAGCCTCACCCCC  
 CTGGGCCTGGCGCTGGTGCTGTGGACAGTGCTTGGGCCCTGCT**TGA**CCCCCAGCGGACACAAGA  
 GCGTGCTCAGCAGCCAGGTGTGTGTACATACGGGGTCTCTCTCCACGCCGCCAAGCCAGCCGG  
 GCGGCCGACCCGTGGGGCAGGCCAGGCCAGGTCTCCTGATGGACGCTGCCGCCCGCCACC  
 CCCATCTCCACCCCATCATGTTTACAGGGTTTCGGCGGCAGCGTTTGTTCCAGAACGCCGCCTC  
 CCACCCAGATCGCGGTATATAGAGATATGCATTTTATTTTACTTGTGTAAAAATATCGGACGA  
 CGTGGAATAAAGAGCTCTTTTCTTAAAAAAA

**FIGURE 382**

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRIF  
LHG NRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPATF  
HGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLHG NR  
ISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAPLRAL  
QYLRLNDNPWVCD CRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCAVATGP  
YHPIWTGRATDEEPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSPPGNGSGPRH  
INDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQAGSGGGGT  
GDSEGS GALPSLTCSLTPLGLALVLWTVLGPC

**Important features:****Signal peptide:**

amino acids 1-26

**Leucine zipper pattern.**

amino acids 135-156

**Glycosaminoglycan attachment site.**

amino acids 436-439

**N-glycosylation site.**

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

**VWFC domain**

amino acids 411-425

**FIGURE 383**

TTCTGTGACCCCTTGAGAAAAGAGTTGGTGGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTCCTGAACTTGTCTGT  
AAGCCCTTGTCCGTAAGCCTTGAACCTACGTTCTTAAATCTATGAAGTCGAGGGACCTTTCGCTGCTTTTGTAGGG  
ACTTCTTTTCTTGTCTTACGCAACATGAGGCTTTTCTTGTGGAACGCGGTCTTGACTCTGTTCTGTCACCTTCTTTGA  
TTGGGGCTTTTATCCCTGAACCAGAAGTGAAAATTGAAGTTCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCA  
AAGGAGGGGATTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTTACTCCACTCACA  
AACATAACAATGGTCAGCCCATTTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGTTGGGACCAGGGCTTGA  
AAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGCTCTGGGCTATGGAAAAGAAGGAAAAGGTA  
AAATTTCCCCCAGAAAGTACACTGATATTTAATATTGATCTCCTGGAGATTGCAAATGGACCAAGATCCCATGAAT  
CATTCCAAGAAATGGATCTTAATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGT  
TTGAAAAACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTGATAAAGAAGATG  
AAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTATAGAGATACATCTACCCTT  
TTAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAAAGAACATTTTATTTTTATACAATGTTCTTTCT  
TGCTTTGTTTTTTTATTTTTATATATTTTTTCTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTT  
CTTTCTGATAAGTTATTGGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTTCACTTTC  
ACAGATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACCACGACATGA  
GACCAGGTTATAGCACAAATTAGCACCCCTATATTTCTGCTTCCCTCTATTTTCTCCAAGTTAGAGGTCAACATTT  
GAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCATGTTATAATGAAATAGTTTATGTGTAAGTGGCTCTG  
AGTCTCTGCTTGAGGACCAGAGGAAAATGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGAT  
GTGCAATGCTGAAGTTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGA  
GGCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAACCCCTATCTCTAC  
TAAAAATACAAAGTAGCCCGGCGTGTTGATGCGTGCTGTAATCCCAGCTACCCAGGAAGGCTGAGGCGGCAGAA  
TCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAGATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAA  
AAGAACACGGTTAATACCATATNAATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGG  
CTCCTAGTGATTGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATGTA  
TCAATATGTTATAGATGAGGTAGAAAAGTTATATTTATATTCAATATTTACTTCTTAAGGCTAGCGGAATATCCTT  
CCTGGTTCCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACATTGTATCATAAGATAAAGTAGTAAACCA  
GTCTACATTTTCCCATTTCTGTCTCATCAAAAACTGAAGTTAGCTGGGTGGTGGCTCATGCCTGTAATCCAG  
CACTTTGGGGGCCAAGGAGGCTGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCT  
TGCTCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAGGCTGA  
GACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCCACTGCACTCCAGCCTGGG  
TGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGACCTACAGCAGCTACTATTGAATAAATACCTA  
TCCTGGATTTT

## **FIGURE 384**

MRLEFLWNAVLTLFVTSLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGS  
LFHSTHKHNNQQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPESTLI  
FNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVEDIFD  
KEDEDKDGFIAREFTYKHDEL

**Important features:**

**Signal peptide:**

amino acids 1-20

**N-glycosylation site.**

amino acids 176-179

**Casein kinase II phosphorylation site.**

amino acids 143-146, 156-159, 178-181 and 200-203

**Endoplasmic reticulum targeting sequence.**

amino acids 208-211

**FKBP-type peptidyl-prolyl cis-trans isomerase**

amino acids 78-114 and 118-131

**EF-hand calcium-binding domain.**

amino acids 191-203, 184-203 and 140-159

**S-100/ICaBP type calcium binding domain**

amino acids 183-203

FIGURE 385

CTCCACACGGTGTCCAGACGCCAGAAATCGCGGCTTCTGGTCTGCTATGGGGTTGCCCTGCTGCTGCTCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTTGAAGGGGACACTGTGTCCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGTGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAATGAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGAACCTCACCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAGTCTTTACTGATCTCTCTGTTCTGTTTCCAGGACCCTGCTGTCTCCTCCCTCCCCTTCTCCACCTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCCCAGGATTOACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGGGGCTGAGGCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACACAGGAACCTCTCCTCACCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGCTGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGGGTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGTGCTGAGCCTTCTGTTCAGCCGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCACGGAGACACAGAGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCCCCCTTCCCAGGCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGAGCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGTGAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAGTCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGATGTTCCAGCCTGACCTAGAAGCGTTTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTGGAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCAGCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGCCTCATGCCAGTGTGCGACCCTGCCTTCTCCTCCACTCCAGACCCACCTTGTCTTCCCTCCCTGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCTGGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCTGTGAAAAACGTGATTCTGGCCCCCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAGGACTCTGAATTCTAACAATGCCAGTGACTGTCGCACTTGAGTTTGAGGGCCAGTGGGCCTGATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCCCAATAGATCTGCTCTGTCTGCGACACCAATCCACGTGGGGACTCCCCTGAGGCCTGCTAAGTCCAGGCCTTGGTCAGGTCAGGTGCACATGTCAGGATAAGCCCAGGACCGGCACAGAAGTGGTTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCTTTTGAAAAAATGATGAAGAAAACCTTGGCTCCTTCCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGAGAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGGATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGAGAAGGGGTGCGGGGTGGTGGTAAAGTAGCACAACTACTATTTTTTTTTCTTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCTGCTGCCAGGCTGGAGTGCAGTGGCAGCATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATTTTTGTACTTTTAGTAGAGATGGGGTTTTACCATTGTTGCCAGGCTGGTCTTGAACCTCCTGACCTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATGCGGGATTACAGGCATGAGCCACTGTGTCTGGCCCTATTTCCCTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAACTTGGAAGATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCATAGTCTCACCAGAGACTATCATTATTTCTGTTTTGTGTACTTCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTTTTTTACAGAGCAATTATCTTGTATATACAACCTTGTATCCTGCTTTTTCCACCTTATCGTTCATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTATAAATAAAATGTTTCATCAGCTGCATAAAAAAAAAAAAAA

**FIGURE 386**

MRLLVLLWGCLLLPGYEALLEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCSG  
TIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDSELLISLFVFP  
GPCCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPGTSQ  
YGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSI PMVRILAPV  
LVLLSLLSAAGLIAFCSHLLLWRKEAQQTETQNEKFWLSRLTAEKEAPSQAPEGDVISM  
PLHTSEEEELGFSKFVSA

**Important features:****Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 248-269

**N-glycosylation site.**

amino acids 96-99

**Fibrinogen beta and gamma chains C-terminal domain.**

amino acids 104-113

**Ig like V-type domain:**

amino acids 13-128

**FIGURE 387**

CGCCGCGGAGCCCATCTGCCCCACAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCGGCACATG  
GCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTCGGAG  
CGCCCGGCCGCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATCGGG  
**ATG**TCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCACACT  
GAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGCTTCCA  
GAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAAGTGGTG  
ATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCGAGTGGC  
TTTGCTTCCAATTTTCTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGCCAGTGAT  
GAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCATGTCATCTTA  
AAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGACAGAAGGAAGT  
GACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTTGTGTATTACTGGCAGCGA  
ATCCGAGAGAAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATTGACTACAACCAC  
CCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTACTCTGGACTGTACCAGTGCACAGCA  
GGCAACGAAGCTGGGAAGGAAAAGCTGTGTGGTGCAGTAACCTGTACAGTATGTACAAAGCATC  
GGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTGATTTTCTCTTTGGTG  
TGGCTGCTAATCCGAAGGAAAAGACAAAAGAAAGATATGAGGAAGAAGAGAGACCTAATGAAATT  
CGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCTCCTCTTCTCAGGCTCT  
CGGAGCTCACGCTCTGGTTCTTCTCCTCCACTCGCTCCACAGCAAATAGTGCCTCACGCAGCCAG  
CGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCACCCAGGCATACAGCCTAGTG  
GGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATGCTAATCTGACCAAAGCAGAA  
ACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAAACGGTCT**TGA**ATTACAATGGAC  
TTGACTCCACGCTTTTCTAGGAGTCAGGGTCTTTGGACTCTTCTCGTCATTGGAGCTCAAGT  
CACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCAGTGAGCATTGCACGGAACAGATT  
CAGATGAGCATTTTTCTTATACAATACCAAACAAGCAAAAGGATGTAAGCTGATTCATCTGTA  
AAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGGAAAGCAGGAGTCCAAATCTATTTGT  
TGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTGAGGTGAATATACCTAAAACCTTTTAAT  
GTGGGATATTTTGTATCAGTGCTTTGATTACAAATTTTCAAGAGGAAATGGGATGCTGTTTGT  
AAATTTTCTATGCATTTCTGCAAACCTTATTGGATTATTAGTTATTTCAGACAGTCAAGCAGAAC  
CCACAGCCTTATTACACCTGTCTACACCATGTACTGAGCTAACCCTTCTAAGAAACTCCAAA  
AAAGGAAACATGTGTCTTCTATTCTGACTTAACTTCATTTGTCATAAGGTTTGGATATTAATT  
TCAAGGGGAGTTGAAATAGTGGGAGATGGAGAAGAGTGAATGAGTTTCTCCCACTCTATACTA  
ATCTCACTATTTGTATTGAGCCCAAAATAACTATGAAAGGAGACAAAAATTTGTGACAAAGGA  
TTGTGAAGAGCTTTCCATCTTCATGATGTTATGAGGATTGTTGACAAACATTAGAAATATATA  
ATGGAGCAATTGTGGATTTCCCCTCAAATCAGATGCCTCTAAGGACTTTCTCTGCTAGATATTT  
CTGGAAGGAGAAAAATACAACATGTCATTTATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAA  
AAGGGATCTAGGAATGCTGAAAGATTACCCAACATACCATTATAGTCTCTTCTTTCTGAGAAA  
ATGTGAAACCAGAATTGCAAGACTGGGTGGACTAGAAAGGGAGATTAGATCAGTTTTTCTCTTA  
ATATGTCAAGGAAGGTAGCCGGGCATGGTGCCAGGCACCTGTAGGAAAAATCCAGCAGGTGGAG  
GTTGCAGTGAGCCGAGATTATGCCATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

**FIGURE 388**

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHQLGLPEKDTLDIEWLLTDNEGNQKV  
ITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHVIL  
KVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRIDYNH  
PGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLIFLLV  
WLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANSASRSQ  
RTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

**Important freatures:****Signal sequence:**

amino acids 1-16

**Transmembrane domain:**

amino acids 232-251



**FIGURE 389**

GCGGCACCTGGAAGATGCGCCCATTTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCTTCG  
CCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCTGT  
CCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCCTCAAAGCTCCAGTCCCCA  
AAAGGCAAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTACTCA  
GCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGGAGAAC  
AGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAAGTGGGAATGTGA  
CAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAGTTTATTC  
AGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCACAAGACTGA  
ATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAGTAAAGAAATCAGGAACATGAAATTCA  
GGTCTAGCTGGGTATTTATTGCAGCAAAAGGCTTGGAAGTCCCTTCCGAAATTCAGAGAGAAA  
AGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGAGATCCAGATAG  
AAGGCTGCATACCCAAAGAACGAAGCTGACACTGCAGGGTCCTGAGTAAATGTGTTCTGTATA  
AACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACAGCCCATATTTGAT  
GAGTATTTTGGGTTTGTGTAACCAATGAACATTTGCTAGTTGTATCAAATCTTGGTACGCA  
GTATTTTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGAACTAAAATGAATGG  
AAATTCTTAAAAAAAAA

## **FIGURE 390**

MRPLAGGLLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQKC  
DHWTCPSPDITYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTATRC  
FDMYEGDNSGPMTKFIQSAAPKSLLFMVTDGSTRLNNDKNAIEALGSKEIRNMKFRSSWV  
FIAAKGLELPSEIQREKINHSDAKNNRYSGWP AEIQIEGCIPKERS

### **Important features:**

#### **Signal sequence.**

amino acids 1-20

#### **N-glycosylation sites.**

amino acids 120-124, 208-212

#### **Glycosaminoglycan attachment site.**

amino acids 80-84

#### **N-myristoylation sites.**

amino acids 81-87, 108-114, 119-125

GGGGGCTTTCTTGGGCTTGCTGCTTGGAAACACCTGCCTCCAAGGACCGGCTCGGAGGGGTGCGCGGGAAGG  
AGGGAAGAAGGAAGGGCGGGGCGGGCCCCCTGCGCCCGCCCCGCGCCTCTGCGCGCCCCGTGTCCGCCCGGGCCC  
AGCCCAGCCCCAGCCCCGCGGGCCGGTACACGCGCAGCCAGCCGGCCGCCTCCCGCGCCAAGCGCGCCGCTCTG  
CTGTGCCCCGCGCCTTGCCCCGCGCAGCTTCTGCGCCCGCAGCCCGCCCCGGCGCCCCCGGTGACCGTGACCCCT  
GCGTGGGCGCGGGGCGGAGAGGCAATGCTCCCCCGGGGACCGCTACCCAGCGCTGGCCCCGTGGTGCTCCTGGC  
AGTGACCCTGGCCGGGTTCGAGAGCCAGGGCGCAGCCCTCGAGACCCTGATTATTACGGGCAGGAGATCTGGAG  
CCGGGAGCCCTACTACGCGCGCCCGAGCCCGAGCTCGAGACCTTCTCTCCGCCGTGCTCGCGGGCCCGGGGA  
GGAGTGGGAGCGCGCCCGCAGGAGCCAGGCCGCCCAAGAGGGCCACCAAGCCCAAGAAAGCTCCCAAGAGGGGA  
GAAGTCGGCTCCGGAGCCGCGCTCCACCAGGTAAACACAGCAACAAAAAGTTATGAGAACCAAGAGCTCTGAGAA  
GGCTGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTGAGAGAGATGGCTGCCACCTCTTGGTCTGGA  
AACCTTAAAAATCACAGACTTCCAGTCCATGCCTCCAGGTGAAGCGCTATAGGCTGGGGGCACATCGAGGGAG  
ACTCAACATCCAGGCGGGCATTAAATGAAAATGATTTTTATGACGGAGCGTGGTGCGCGGGAAGAAATGACCTCCA  
GCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCATCACTCAAGGGGAGGAACCTCCTCTG  
GCTGAGTGACTGGGTGACATCCTATAAGGTGATGGTGAGCAATGACAGCCACACGTGGGTCACTGTTAAGAATGG  
ATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCTACCCGTCCCATGGT  
GGCCCGCTACATCCGCATAAACCTCAGTCTGGTTGTATAATGGGAGCATCTGCATGAGAATGGAGATCCTGGG  
GTGCCACTGCCAGATCCTAATAATTATTATCACCGCGGAACGAGATGACCACCTGATGACCTGGATTTTAA  
GCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTAATGAAATGTGTCCCAATGATCACCAGAAATTA  
CAACATTGGAAGAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATCACCTGGGGAGCATGAAGT  
CGGTGAGCCCGAGTTCCACTACATCGCGGGGGGCCACGGCAATGAGGTGCTGGGCCGGGAGCTGCTGCTGCTGCT  
GGTGCAAGTTCGTGTGTGCTCAGGAGTACTTGGCCCGGAATGCGCGCATGCTCCACCTGGTGGAGGAGACGCGGATTCA  
GTCCTCCCCCTCCCTCAACCCCGATGGCTACGAGAGGCCCTACGAGGGGGCTCGGAGCTGGGAGGCTGGTCCCT  
GGGACGCTGGACCCACGATGGAATTGACATCAACAACAACCTTCTGATTAAACACCGTCTCTGGGAGGCAGA  
GGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCCTGAGTGGTTTCTGTGCGAAATGC  
CACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAAAAATCCCTTTTGTGCTGGCGGCAACCTGCA  
GGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGCGGTCCCCCTGGAAGACGCGAGGAACACACCCCCAC  
CCCCGATGACCAGTGTTCCGCTGGCTGGCCCTACTCCTATGCCTCCACACACCGCCTCATGACAGACGCCCCGGAG  
GAGGTGTGGCCACACGGGAGGACTTCCAGAAGGAGGAGGCATGTCAATGGGGCTCCTGGCACACCGTCTGGTGG  
AAGTCTGAACGATTTTCACTACCTTCATACAACTGCTTGAAGTTCATGCTACGTGGCTGTGATAAATACCC  
ACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGGAATCTCTGATCGTGTTCATGGAGCAGGTTTCATCG  
TGGCATTAAAGGCTTGGTGAGAGATTACATGGAAAAGGAATCCCAAACGCCATTATCTCCGTAGAAGGCATTAA  
CCATGACATCCGAACAGCCAAACGATGGGGATTACTGGCGCCTCCTGAACCCCTGGAGAGTATGTGGTCACAGCAA  
GGCCGAAGGTTTCACTGCATCCACCAAGAATGTATGGTTGGCTATGACATGGGGGCCACAAGGTGTGACTTTCAC  
ACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGAGAAGTTTGGGAAGCAGCCCGTCAGCCTGCCAGC  
CAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGTGGGTGACACCTCCTGGGCCCTTGAGACTCGTCTGGG  
ACCCATGCAAAATTAAACCAACCTGGTAGTAGCTCCATAGTGGACTCACTCACTGTTGTTTCTCTGTAATTCAAG  
AAGTGCCTGGAAGAGAGGGTGCATTTGTGAGGCAGGTCCCAAAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTT  
CTTTGTTCCCATTTTATCCAAATAACTTGGACAGAGCAGCAGAGAAAAAGCTGATGGGAGTGAGAGAACTCAGCAAG  
CCAACCTGGGAATCAGAGAGAGAAGGAGAAGGGGAGCCTGTCCGTTGAGAGCCTCTGGCTGCATGAAAAAGG  
ATTCTGGTGCTTCCCTGTTTGTGCTGGCAGCAAGGGTCCACGTGACATTTGCAATTTGCACAGCTAAAATTGCG  
CATTTCCCAGCTGGGCTGTCCCAAAATGTTACCATTGAGATGCTCAGGCGCTCCTAAGAGAATCCACCCTCTC  
TGGCCCTGGGACATTGCAAGCTGCTACAAATAAATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACA  
TCAGTGAGCCTCTTGAATCTGTTTAGTCTCCTTTTTCAACAAAGGAGTGTGTTTCAGAAAAGGAGAGAGAGGCTGA  
GATCATTCAGGAGTTTGTGGGCGAGCAAGCATGGAGCTTCTTGCAAAATCTGGGTCCATAAACAACCCCCAAA  
TCCCTGCTGATCAGTAGCCCTGGAGGTTCCCCAGGTAGGAGAGCCAGAGTGGCAGCCTTCTGAAGGGCCA  
GAAAATTTAGCCTGGATCTCCTCTTTTACCTGCTAGGACTGGAAGAGAGCCAGAAGTGGGCTGGCCTGAAGCCCTC  
TCTCTGCTTGAGGTATTGCCCTGTGTGGAATTGAGTGCTCATGGGTGGCCTCATATCAGCCTGGGAGTTATTTT  
TTGATATGTAGAATGCCAGATCTTCCAGATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCAT  
CAGTTTGGGAAGAATTATTGAATTATCTTGCAAGAAAAAGTATGTCTCACTTTTGTTAATGTTGCTGCCTCAT  
TGACCTGGGAAAAATGAAAAAATAAAGCAAATGGTAAGACCCTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAA

**FIGURE 392**

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPPPLPA  
GPGEWEERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHSVR  
VAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCAGR  
DLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIFEGNS  
EKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTTTDDLD  
FKHHNYKEMRQLMKVNVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEFHYIAGA  
HGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGGSELGGWS  
LGRWTHDGIDINNNFIDLNTLLWEAEDRQNVPRKVPNHYYIAIPEWFLSENATVAAETRAVIAW  
MEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSASTHRLMTDARR  
RVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHESQLPEEWENNR  
ESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLNPGYVVTAKA  
EGFTASTKNCMVGYDMGATRCDFTLSTNMARIREIMEKFGKQPVSLPARRLKLRGRKRRQRG

GTCCACACATCTCTGCTCAACTGGGTGAGTCCCTCTTAGACCAGTCTTGTCCATCATTTGCTGAAGTGGACCAAC  
TAGTTCCCCAGTAGGGGGTCTCCCCTGGCAATTCTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA  
TGGCCTTGCCTTGGGGTCTGCTTGTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG  
AAGGAGCAGGGGGTCTGATCAAGCCATCCAGGAACACTGGAGGACTTGTCCAGCCTTGAAGAAGCTCTAGTGGTT  
TCTGAATCTAGCCCACTTGGCGGTAAGC**ATG**ATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGTTGG  
CTACTTATTTCTTTTAGGGGATTGTGAGGAGTACCACCTCTCACGGTGAATACCAAGTGTACAGGAAGTGCC  
ATCTGGTACAGTGATCGGGAAGCTGTCCCAGGAAGCTGGGCGGGAGGAGAGGCGGAGGCAAGCTGGGGCCGCTT  
CCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCATTACAGGTGGACTCTGAGGAAGGCTTGTCTCAGCACAGGCAGGCG  
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCTGGTTTCCCTTTGATGTGCTTGCCACAGGGGATT  
GGCTCTGATCCATGTGGAGATCCAAGTGTGGACATCAATGACCACCAGGCCACGGTTTCCCAAAGGCGAGCAGG  
GCTGGAATCTCTGAGAGCGCTCTGTGCGAACCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACAGGCC  
TAAACCCCTGCACACCTACACTCTGTCTCCAGTGAGCACTTTGCCCTGGATGTCAATTTGGGCCCTGATGAGAC  
CAAACATGCAGAACTCATAGTGGTGAAGGAGCTGGACAGGGAATCCATTCATTTTTTGATCTGGTGTAACTGC  
CTATGACAATGGGAACCCCCCAAGTCAAGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAG  
CCCTGCGTTTGTCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAACT  
GACCCGCCACACCTTGACCAAGGCCCCAATGGGGAGGTGGAGTTCTTCTCAGTAAGCACATGCCCTCCAGAGGT  
GCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCACTTCGCTCGACCTTAGACTATGAAAAGAACCC  
TGCCTACGAGGTGGATGTTCAAGCAAGGAGCTGGGTCCCAATCTATCCAGCCCATTGCAAAGTCTCATCAA  
GGTTCTGGATGTCAATGACAACATCCAAGCATCCACGTACATGGGCCCTCCAGCCATCACTGGTGTCTGAAAG  
TCTTCCCAAGGACAGTTTTATTGCTCTTGTCTATGGCAGATGACTTGGATTGAGGACACAATGGTTTGGTCCACTG  
CTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC  
CACACTGGACAGAGAGCAGTGGCCCCAATATACCTCCTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC  
AGCCAAGAAACAGCTCAGCACTCAGATCAGTACATCAACGACAATGCACCTGTGTTTGAAGAAAGCAGGTATGA  
AGTCTCCACGCGGGAAACAACTTACCTCTCTTCACTTCACTTACCAAGCTCATGATGACAGCTTGGGCAT  
TAATGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA  
GGTCACTGCTCAGAGGTCACTGAACATATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG  
GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTGAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT  
CAGCCTGTGCTCAGCATGGAAGCCAGCCTCTCCGTGCTTGTGAATGCCCTCCACAGGCCACCTGCTGGTGCC  
CATCGAGACTCCCAATGGCTTGGGCCAGGGCACTGACACCTTCACTGGCCACTCAGACTCCCGGCCATT  
CCTTTTGACAACCATTTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCTCTACAGCATCCGCAATGG  
AAATGAAGCCCACCTCTTCATCTCAACCTCATACGGGGCAGCTGTTGCTCAATGTACCAATGCCAGCAGCCT  
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCTTACAGACCCGAGCCCTGTT  
GAGGGTCATGTTTGTCTACCAAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCCTTGAGCATGTGAT  
GCTCAGCGTGATCTGCCGTGACTGTGTCCTTGGGCACTTTCGGTTGATCTGGCTTGTCTATGTCATGTGCCG  
GACAGAAAAGAAGGACAACAGGGCCCTACAACATGTGCGGAGGCGGATCCACTACCGCCAGCAGCCCAAGAGGCC  
CCAGAAACACATTGAGAAGGCAGACATCCACCTCGTGCTGTGCTCAGGGGTGAGGCAGGTGAGCCTTGTGAAGT  
CGGGCAGTCCCACAAAGATGTGGACAAGGAGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTT  
CCACCTCACCCCGACCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGT  
TGCTCAAGACAGCGTCAACCTCCTTTCAACCATCCAGGCAGAGGAATGCCCTCCCGGGGAACCTGAACCTTCC  
CGAGCCCCAGCCTGCCACAGGCCAGGCTTCCAGGCTCTGAAGGTTGCAGGACGCCCCACAGGAGGCTGGC  
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACCAGCCTCCTCTGCAACCTGAGACGCGAGCAGACATCT  
CAATGGCAAAGTGTCCCCTGAGAAAGAATCAGGGCCCCGTGAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC  
TGCCTTTCGCGGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTTACGAAATCTCCAGCTGCT  
GTCCTTGTGTCATCAGGGCCAATTCCAGCCCAAACCAAACACCGAGGAAATAAGTACTTGGCCAAGCCAGGAGG  
CAGCAGGATGCAATCCAGACACAGTGGCCCCAAGTGCAAGGCTGGAGGCGAGACGACCCAGAACAGGAGGA  
AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCACTGCTAGAAGAGAGCTGTCAAGTCTGTCTGGACCC  
CAGCACAGGTCTGGCCCTGGACCGCTGAGCGCCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCTCAC  
CACCACCTACCTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTCCGG  
CAAGGCAGAGGCACCAGAGCTGAGCCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTCTCGGAGATGAGCTCACT  
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCGCGCTCCGAGGCGCTGCGGGCGGCTCTCGGT  
TGCGGGAGAGCAGCTCACTTTAGACTTGGCCACCATGACGCTCAGGCTCAGGCTGAAAGTGAACAGGGGACCCAGGTGG  
AAAGACGGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGGAGGTGCTG**TGA**ACATACTCAGACGCT  
CTGGATCCAAGAACCAGGGGCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAAATCTTGTAACTCACTAGCTAG  
CGCGGGCTGAGAACTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC  
TGACCAAAGCAGCCCCCTGTGAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGT  
TCTTGCCAAAACATATGTGGAGCACAAGGGTCAGTCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG  
AAAGGTTGGCCTTCTTGGTAGCAGGAGTCAGGGGCTGTACCCTGGGGGTGCCAGGAAATGCTCTCTGACCTAT  
CAATAAAGGAAAAGCAGTAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 394**

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQAG  
 AAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQVL  
 DINDHQPRFPKGEELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIVGPD  
 ETKHAELIVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNDNSPAFAESSLALE  
 IQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPLDYEKN  
 PAYEVDVQARDLGNPIPAHCKVLIKVLVDNDNIPSIHVTWASQPSLVSEALPKDSFIALVMA  
 DDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQDQGLQPLS  
 AKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVSYRIQDSPV  
 AHLVAIDSNTGEVTAQRSLNVEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDANDNAPEVVQ  
 PVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLTTIVARDADSGA  
 NGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGSPPLQTRALLRV  
 MFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEKKDNRAYNCREAE  
 STYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEAGWDPCLOAPFHLT  
 PTLYRTRLRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLEPPQPATGQPRSRPLKV  
 AGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQILRSLVRLSVAFAE  
 RNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGSRSAPDTPDGPSARAGG  
 QTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPDPAWMARLSLPLTTNYRD  
 NVISPDAATEEPRTFQTFGKAEPALSPTGTRLASTFVSEMSSLLEMLLEQRSSMPVEAASE  
 ALRRLSVCGRTLSLDLATSAAAGMKVQGDPPGGKTGTGEGKSRGSSSSSRCL

**Important features:****Signal peptide:**

amino acids 1-13

**Transmembrane domain:**

amino acids 719-739

**N-glycosylation site.**

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

**Cadherins extracellular repeated domain signature.**

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

**FIGURE 395**

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAGG  
CTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAATC  
AGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACC**ATG**GGACGCCCCCGACCTCGTGCG  
GCCAAGACGTGGATGTTCTTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGCACAG  
GAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGGCCTTG  
TTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTTACAGCT  
GCCCCTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGAT  
GGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGAT  
GTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCC  
AAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCA  
GGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCTTGACACTCTCAACTGTGCAGAAGTA  
AAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATGGCATGGTC  
TGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGT  
GATGGTGCCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCT  
GGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATCATAGGCAGCAAGGGC  
**TGA**TTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACCTCTCTGGTTC

## **FIGURE 396**

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVLV  
GGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLLQL  
RDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPTLNCAEVKIFPQKKCEDAYP  
GQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWI  
KKIIGSKG

### **Important Features:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 51-71

#### **N-glycosylation site.**

amino acids 110-113

#### **Serine proteases, trypsin family, histidine active site.**

amino acids 69-74 and 207-217

#### **Tyrosine kinase phosphorylation site.**

amino acids 182-188

#### **Kringle domain proteins motif**

amino acids 205-217



**FIGURE 397**

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCCATGTCGGACCTGCTAC  
TACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGGGT  
ACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCACTG  
TGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGCTGCA  
GCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCCCCCTG  
ATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCCCTGAGC  
TCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGCCATGTGG  
TGACAGCCACCTTCCCCTACACCACCATTCTGTCCATCTGGCTGGCTACCCGCCGTGTCCATC  
CTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGGAGATCTACC  
AGGAAGACCAGATCCATTTTCATGTGCCACTGGCACGGCAGGGAGACTTCTATGTGCCTGAGA  
TGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCAGGTGGATGGCA  
CAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCCCTGGCAGCCGGG  
AGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGATGACGGTGACACCC  
GCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGAGGAGCTGGACTTGG  
AGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGCCCCTGGGGACTACCA  
AGTGGCTCTGGGAGCCCCTGCCCCTGAGAAGGGCAAGGAGTAAACCCATGGCCTGCACCCTCC  
TGCAGTGCAGTTGCTGAGGAACCTGAGCAGACTCTCCAGCAGACTCTCCAGCCCTCTTCCTCCT  
TCCTCTGGGGGAGGAGGGGTTCCCTGAGGGACCTGACTTCCCCTGCTCCAGGCCTCTTGCTAAG  
CCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCAGGGACTATTTTCTGCACCA  
GCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTTTCAGACTCACAGTGGAGCTTCCAGGACC  
CAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAAAAAAAAAAAAAAA

**FIGURE 398**

MSDLLLLGLIGGLTLLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGlyGETGRL  
FTESCSISPKLRsIAVYYDNPHMVPPDKRCaVGSILSEGEESpSPeLIDLYQKFgFKVfSfP  
APSHVVTATFPYTTILSIWLATRrVHPALDtyIKERKLCAYPrLEIyQEDQIHfMCPLARQGD  
FYVPEMKETEWKWRGLVEaIDtQVDGTGADtMSDTSSVSLEvSPGSREtSAATLSPGASSRGW  
DDGDTRSEHSySESGASGSSFEELDLEGEgPLGESRLDPGTEPLGTTKWLWEPTAPEKGE

**FIGURE 399**

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGGC  
 TCCGTGCCGCCAAGTTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCGAG  
 AGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTGTTTTTTAAACT  
 TCTGTTTCTTGGGAGGGGGTGTGGCGGGGCAGG**ATG**AGCAACTCCGTTCCTCTGCTCTGTTTC  
 TGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGACGGCTG  
 GAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGGTTTAAC  
 CTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTGCGCCACAGCCAGCCC  
 TTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTTCATCATTCACGGATGGACGATG  
 AGCGGTATCTTTGAAAACCTGGCTGCACAAACTCGTGTCAGCCCTGCACACAAGAGAGAAAGAC  
 GCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGATGCGGTCAAT  
 AATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGGAGAAGGACGAT  
 TTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTGGCCGGGTATGCA  
 GGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGCCGGGGCCCATGTTT  
 GAAGGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTGTGGATGTCCTCCAC  
 ACCTACACGCGTTCCTTCGGCTTGAGCATTGGTATTGAGATGCCTGTGGGCCACATTGACATC  
 TACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGTCTTGGGATCAATTGCA  
 TATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCGTCCACCTCTTTGTTGAC  
 TCTCTGGTGAATCAGGACAAGCCGAGTTTTTGCTTCCAGTGCAGTCACTCCAATCGCTTCAAA  
 AAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCATTGGCTACAATGCCAAGAAA  
 ATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGGCAGGCATGCCTTTCAGAGGT  
 AACCTTCAGTCCCTGGAGTGTCCC**TGA**GGAAGGCCCTTAATACCTCCTTCTTAATACCATGCT  
 GCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCACAAATCCAATCAAATCGTTGCAA  
 ATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTTACAAAATAAACAGTGTGGACCCC  
 TAATAA

**FIGURE 400**

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHEG  
CYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWLPL  
AHQLYTDAVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTVGRI  
TGLDPAGPMFEGADIAHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDFQPGC  
GLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLQNQDKPSFAFQCTDSNRFKKGICLSCRKNR  
CNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

**Important features:****Signal peptide:**

amino acids 1-16

**Lipases, serine active site.**

amino acids 163-172

**N-glycosylation sites.**

amino acids 80-83 and 136-139

**FIGURE 401**

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCTG  
CCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCTGT  
CAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGGCTC  
AACTCTCCTGCACGCTCAGCCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGGTACC  
AGCAGCGGGCAGGCAGTGCCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCACCACC  
GGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCACAATGCCTGTGTCC  
TCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGCTACGGCT  
TTAGTCCC**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCCCTCTGCCTCCCATTTCTGCCCCCTGA  
CCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGGTAAATAATA  
TTCAACATGTCAACAAC

402/550

**FIGURE 402**

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAGS  
APRYLLYYRSEEDHHRPADIPDRFSAKDEAHNACVLTISPVQPEDDADYYCSVGYGFSF

402/550

**FIGURE 403**

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGAAAGCAGCGAGTTG  
GCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGTGGCTCACAACAAGATGCTCAAGGTGTCAGC  
CGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCAGTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGG  
GCGGTCCGACGGCGGTAATTTTCTGGATGATAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGG  
ACAGTGGAACAAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCTTCGATCA  
GGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTATGCATTGCTCAAGATTC  
TCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATGAAAGAAGCAGGAGTAGACCATAGGCAGTG  
GAGGGGTCCCATATTATCCACCTGCAAGCAGTGCCAGTGGTCTATCCAGCCCTGTTTGTGGTTCAGATGGTCA  
TACCTACTCTTTTCAGTGCAAACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGG  
ACATTGCCCATGTCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCTGGAGTT  
CAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAGTCAAAACAAGAAGACAAA  
AACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATCTTGCCAATTTGCAAGGACTCACTTGGCTGGAT  
GTTTAACAGACTTGATACAACTATGACCTGCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAA  
TGAACAGTGTACCAAGGCATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTG  
CTACTGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCAAGGGGTAAA  
GAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGCCAACACAATGTCATGGCAGTGT  
TGGACAGTGTGGTGTGTTGACAGATATGGAAATGAAGTCATGGGATCCAGAATAAATGGTGTTCAGATTGTGC  
TATAGATTTTGGAGATCTCCGGAGATTTTGTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGA  
CGATATTATGAATGATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGA  
CCATGATGTATACATTGATTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTACAAAAATGATAG  
CCTATTTAAAATTATCTTCTTCCCAATAACAAAATGATTCTAAACCTCACATATATTTTGTATAATTATTTGAA  
AAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAATAAGAATCATTTGCTTTGAGTTTTTATATTCTTACACA  
AAAAGAAAATACATATGCAGTCTAGTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTACGAGA  
ACAAACTTTGTAAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAGAT  
AATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGGAAAAATATGCATGCT  
TTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAGGATAACAGAGAGATACCACATGACTCCA  
AAAAAAAAAAAAA

**FIGURE 404**

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFRD  
EVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEAGV  
DHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPGPSDK  
PTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLRPERSRFDTSILPICKDS  
LGWMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQRQQDP  
PCQTELSNIQKRQGVKKLLGQYIPLCEDDGYYKPTQCHGSVGQCWCVDTRYGNEVMGSRINGVA  
DCAIDFEISGDFASGDFHEWTDDEDDDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

**Important features:****Signal peptide:**

amino acids 1-16

**Leucine zipper pattern.**

amino acids 246-267

**N-myristoylation sites.**

amino acids 357-362, 371-376 and 376-381

**Thyroglobulin type-1 repeat proteins**

amino acids 353-365 and 339-352



[illegible]

**FIGURE 406**

MTPQSLLQTTFLFLLSLLFLVQGAHGRGHREDFRCSQRNQTHRSSLHYKPTPDLRISIENSEE  
 ALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLSDKASSLLCFQHQE  
 ESLAQGPPLLATSVTSWWSPQNISLPSAASFSTFSFHSPHTAAHNASVDMCELKRDQLLSQF  
 LKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQDLHIH  
 SRQEEEQSEIMEYSVLLPRTLQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGEKVLGI  
 VVQNTKVANLTEPVVLTQFQHQLPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRETQTSCE  
 CNHLTYFAVLMVSSVEVDVAVHKHYLSLLSYVGCVVVSALACLVTIAAYLCSRVPPLPCRKRPRDY  
 TIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLEGYNLYRLV  
 VEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIYPSMCWIRDS  
 LVSYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLGLPWALIFFSF  
 ASGTFQLVVLYLEFSIITSFQGFLIFIWWSMRLQARGGPSPLKSNSDSARLPISSGSTSSSRI

**Important features:****Signal peptide:**

amino acids 1-25

**Putative transmembrane domains:**

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590  
 and 634-657

**Microbodies C-terminal targeting signal.**

amino acids 691-693

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 198-201 and 370-373

**N-glycosylation sites.**

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327  
 and 341-344

**G-protein coupled receptors family 2 proteins**

amino acids 475-504

**FIGURE 407**

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGAG  
CCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTCGC  
TCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCTGGA  
CTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAAGCCC  
TGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAAAGCTC  
AGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAGCAGGGC  
TCTCAGAAGGCGGTGGTGCCAGCTGGGATC**ATG**TTGTTGGCCCTGGTCTGTCTGCTCAGCTG  
CCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTGCTACATGA  
CTTCGGGCTGGACGGATAACGGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGCTTATTTAC  
AAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACAACGGGATCTT  
CCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAACGTGTGCCGGAT  
GTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGCCATGAAGATAAC  
CCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCCTGGAGGCATCACTGCCAGGGAAAAGACCT  
CACTGAATGGGTGGATGGCTGTGACTT**CTAG**GATGGACGGAACCATGCACAGCAGGCTGGGAA  
ATGTGGTTTGGTTCCTGACCTAGGCTTGGAAGACAAGCCAGCGAATAAAGGATGGTTGAACG  
TGAAA

## **FIGURE 408**

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALDY  
EADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYSDDLNPNLKDTVICAMKITQEPQGLGYWE  
AWRHHCQGKDLTEWVDGCDF

### **Important features:**

#### **Signal peptide:**

amino acids 1-18

#### **N-myristoylation site.**

amino acids 67-72

#### **Homologous region to Alpha-lactalbumin / lysozyme C proteins.**

amino acids 34-58 (catalytic domain), 111-132 and 66-107

**FIGURE 409**

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC  
 CTCTTTTCAGCCCGGGATCGCCCCAGCAGGG~~ATG~~GGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC  
 GCTCTGCCTCCGGTGCTGCTGCGCTGGGGCGGCCGGCTTACACCTTCCCTCGATAGCGACTTCACCTTTACCCTT  
 CCGCGCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCCGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTTA  
 GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCAAAACCTTAGTTTTTTGAACAAAGAAAA  
 TCAGATGGAGTTACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTACAGCACCATT  
 TCTGAGAAGGTGATTTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG  
 AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC  
 AGACTAAGCAAAAGTGGGCACATACAAATCTGCTTAGAGCATTGGAAGCTCGTGATCGAAACATACAAGAAAGC  
 AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGTGAGCCATTCAAGTTTAT  
 ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACT~~TAA~~AACTCCAACTAGAGTACGTAACATTGAAA  
 AATGAGGCATAAAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTTGAGATATA  
 AAAGTAGGAAACAGGTATAATTTTAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG  
 TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTAACTGAATGAAGCCATATTAATAACTGCAT  
 TTTCTTAACCTTTGAAAAATTTTGCAAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC  
 AGTCTAGTTTTTAAACAGGTTCTATTACCCAGAACCTTTTTTGTAATGCGGCAGTTACAAATTAACGTGGAAGTTT  
 TCAGTTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCCAA  
 CTTTTCTCTATTTACATATGCATCTCTCCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG  
 AGATTTTTTATAACCAAATACATTTTCAGTGTAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC  
 CCAAAGCTGACATTTTCACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG  
 AAATGAAGAATATAGTTTAAAAGCTTCCCTCCATAGGGACACATTTTCTTAACCCTTAACTAAAGTGTAGGA  
 TTTTAAAATTAATGTGAGGTAAAATAAGTTTATTTTTAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA  
 TAATCATGTTATGTTAATTTTAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA  
 TTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA  
 AGAATTTAGAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACATATAAATAAATATCTAGA  
 ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAAACAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG  
 AAACCTTTGGCTGTAGGTTTTTATTTTCTACAAGAATTCTGGTTTGAATTAATTTTTTGTAAGCAGGTACATTTTAT  
 AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTTATAAAAATTTTTATTATAACAACCTTTTAT  
 TAAAATGGCCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTTAA  
 CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC  
 TCAAGTACTAGTAATTTAACTTCATCATGAATGAACATATAATTTTAAAGTTATGCCATTTATAACGTTGTTTAT  
 GACTACATTTGTGAGTTAGAAACAACTTAAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT  
 CTTGATGAGCAATAATGATAACCAGAGAGTGATTTCAATTTACACTCATAGTAGTATAAAAAAGAGATACATTTCCC  
 TCTTAGGCCCTTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT  
 ATGATCAATTACCTTAATTTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA  
 AGGTCAATAAGATCCTTGCCATGAATACCCCTCCCTTTTGCGCTGTTAAATTTGCAATGAGAAGCAAATTTACA  
 GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAACTGTGATTAAGAATTTCTA  
 CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCT  
 ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTACAGTAACCTCTTACCATATAAAAAACGATAATTGCTT  
 TATTTGGAAAAGAATTTAGGAATACTAAGGACAATTTATTTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG  
 CATAACCAAAAAAGCAAACTTGTAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT  
 CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATTGAGTAAGCTGTTTATCACTT  
 AACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTT  
 CATAATGTAGCAGTTACCGTGTTACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT  
 GTTATGCTGTTCTTTTCATGTGAATGTCAAGACATGGAGGGTGTGTAATTTTATGGTAAAATTAATCCTTCTTA  
 CACATAATGGTGTCTTAAATTTGACAAAAAATGAGCACTTACAATTGTATGTCTCCTCAAATGAAGATTCTTTAT  
 GTGAAATTTTAAAGACATTGATTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG  
 CTCAAACTGCTTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA  
 AATTATCAAAGGAAAA

## **FIGURE 410**

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFITLPAGQKECFYQPMPLKASLEIEYQ  
VLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFELIL  
DNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRNIQE  
SNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 195-217

#### **N-myristoylation site.**

amino acids 43-48

#### **Tyrosine kinase phosphorylation site.**

amino acids 55-62

CCACGCTGAGGAGCCCTGTGCTCAAGACACGGTCACTGGTCTGAGAAACTTCCCAGGGGACCGCATTCAGAGTCA  
GTGACTCTGTGAAGCACCACCATCTACCTCTTGCCACGTTCCCACGGGTGGGGGAAAGATGGTGGGGACCAAG  
GCCTGGGTGTTCTCCTCTGCTGCTGGAAGTACATCTGTGTGGGGAGACAGCATGCTCACCCAGTCACTA  
AGAAGAGTCCAGCCTGGGAAGAAGAACCCAGCATCTTTGCCAAGCCTGCCGACACCCTGGAGAGCCCTGGTGAG  
TGACAAACATGGTTCAACATCGACTACCCAGCGGGAGGGGACATGATGAGCGGCTGGACGCCATTCGCTCTCTAC  
TATGGGACCGTGATGTGCCGTCCTCGGCTAGAGGCTCGGACCATGACTGGACACCTGCGGGCAGCACT  
GGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTTCTGGTGCTCAACAGGGAGCAGCGGCCCTGGCCAGAACTGC  
TCTAATTACACCGTAGCTCTCCTCTGCCCCACAGGATCCCTGCGCCGAGACACAGAGCGCATCTGGAGCCCATGG  
TCTCCTCTGGAGCAAGTGCTCAGCTGCCGTGTGGTGACAGTGGGTCCAGACTCGCACAGCATTTGCTTGGCAGAG  
ATGGTGTCGCTGTGCAGTGAGGCCAGCGAAGAGGGTTCAGCACTGCATGGGCCAGGACTGTACAGCCTGTGACCT  
ACCTGCCCAATGGGCGAGGTAATGCTGACGTGATGCTGATGCTGATGCTGAGGACCTTCATGCTTCATGGGGCTGT  
TCCCTTCCCAGAGTGGCCAGCCTCAGGGCTGCTATCTACCTCTGACCAAGACGCCGAAGCTGCTGACCCAG  
ACAGACAGTGATGGGAGATTCCGAATCCCTGGCTGTGTGCCCTGATGGCAAAGCATCCTGAAGATCACAAAGGTC  
AAGTTTGTGCCCCATTTGTATCAACAATGCCAAGACTAGCTGAAAGGACGCCACCATCAAGGCGAGAGTTGTGTGAGG  
CGAGAGATCCATACATGGTGATGAACCTTGAGACAAAGACAGGAGAGCTGGCGAGAGCGTGCTCTGTGCTGT  
AAGGCCACAGGGAAGCCAGGCCAGCAAGTATTTTTGGTATCATAATGACACATATTGCTGGATCCTTCCCTCTAC  
AAGCATGAGAGCAAGCTGGTGCTGAGGAACTCGCAGCAGCAGCCAGGCTGGGGAGTACTTTTGAAGGCCAGAGT  
GATGCTGGGGCTGTGAAGTCCAAGGTTGCCAGCTGATTGTACAGCATCTGATGAGACTCCTGCAACCCAGTT  
CCTGAGAGCTATCTTATCCGGCTGCCCCATGATTGCTTTCAGAATGCCACCAACTCCTTCTACTATGACGTGGGA  
CGCTGCCCTGTTAAGACTTGTGCAGGGCAGCAGGATAATGGGATCAGGTGCCGTGATGCTGTGCAGAACTGCTGT  
GGCATCTCAAGACAGAGAAAGGGAGATCCAGTGCAGTGGCTACACGCTACCCACCAAGGTGCCAAGGAGTGC  
AGCTGCCAGCGGTGACGGAAACTCGGAGCATCTGCGGGGCCGTGTGAGTGCTGCTGACAATGGGGAGCCCATG  
CGCTTTGGCCATGTGTACATGGGGAAACAGCCGTGTAAGCATGACTGGCTACAAGGCGCATTTTCACTCCATGCT  
CCCCAGGACACTGAGAGCTGGTGCTCACAATTTGTGGACAGGCTGCAGAAGTTTGTCAACACCACCAAGTGTCTA  
CCTTTCAACAAGAAGGGGAGTCCCGTGTTCCATGAATCAAGATGCTTCGTCGGAAGAGGCCCATCACTTTGGAA  
GCCATGGAGACCAACATCTATCCCCCTGGGGGAAGTGGTTGGTGAAGACCCCATGGCTGAACCTGGAGATTCCATCC  
AGGAGTTTCTACAGGCAGAATTGGGGAGCCCTACATAGGAAAGTGAAGGCCAGTGACCTTCTGGATCCCCGG  
AATATTTCCACAGCCACAGCTGCCCCAGACTGACCTGAACCTTCATCAATGACGAAGGAGACCTTTCCCCCTCGG  
ACGTATGGCATGTTCTCTGTGGACTTCAGAGATGAGGTCACTTCAGAGCCACTTAATGCTGCGCAAGTGAAGGT  
CACTTGACTCGACCCAGGTCAAGATGCCAGAGCACATATCCACAGTGAACACTTGGTCACTCAATCCAGACACA  
GGGCTGTGGGAGGGAAGGTGATTCAAATTTGAAATCAAAGGAGGAACAAAGAGAGAAGACAGAACCTTCCCTG  
GTGGGCACTCGAGATTCTGTGAGAGGAGGCTCTTTAACTGGATGTTTCTGAAAGCAGGCGGTGCTTTGTTAAG  
GTGAGGGCTTACCGGAGTGAGAGTTCTTGCTTAGTGAGCAGATCCAGGGGGTTGTGATCTCCGTGATTAACCTG  
GAGCCTAGAAGTGGCTCTTGTCCAACCTTAGGGCTGGGGCCGCTTTGACAGTGTCTATCAGAGGCCCAACGG  
GCTGTGTGCTTCCGCTTGATGATGACCACTGCCCTGATGCTTCTGCTTCTGGCAAGCCTGGCTGGG  
GAGGAAGTGCAGAGCTGGAGTCTTCTCTCAAATTCACCCCAATGCAATTTGGCGTCCCTCAGCCCTATCTCAAC  
AAGCTCAACTACCTCCGACGGACCATGAGGATCCACGGGTTAAAGACAGCTTTCCAGATTAGCATGGCCAAAG  
CCAAGGCCCACTCAGCTGAGGAGCAATGGGCCCATCTATGCTTTGAGAAGCTCCGGGCGATGTGAAGAGGCA  
CCACCCAGTGCAGCCCACTTCCGGTTCTACCAGATTGAGGGGGATCGATATGACTACAACACAGTCCCCCTTCAAC  
GAAGATGACCTATGAGCTGGACTGAAGACTATCTGGCATGGTGCCAAAGCCGATGGAATTCAGGGCCCTGCTAT  
ATCAAGGTGAAGATTGTGGGGCCACTGGAAGTGAAATGTGCGATCCCGCAACATGGGGGCACTCATCGGCGGACA  
GTGGGGAAGCTGTATGGAATCCGAGATGTGAGGAGCACTCGGGACAGGGACCAGCCCAATGTCTCAGCTGCCGTGT  
CTGAGTTCAAGTGCAGTGGGATGCTCTATGATCAGGACCGTGTGGACCGCACCTTGGTGAAGGTATCTCCCCGAG  
GGCAGTGCCCTCGAGCCAGTGTGAACCCCATGCTGCATGAGTCACTGTTCAACCACCTTGCCACTTGAGATCAAC  
AACGACACCAGTGAGTACACCATGCTGGCACCCCTTGGACCCACTGGGCCACAACATATGGCATCTACACTGTCAC  
GACCAAGGACCTCGCAGCGGCCAAGGAGATCGCGCTCGGCCGGTGCTTTGATGGCACATCCGATGGCTCTCCAGA  
ATCATGAAGAGCAATGTGGGAGTAGCCCTACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGATGCCCTTC  
CAGTACCTCCAAAGCACCCAGCCAGTCCCCCTGCTGCAGGCACCTGCCAAGGAAGAGTGCCCTCGAGGAGGCAG  
CAGCGAGGACAGCGGGGTGGCCAGGCGAGGTTGGAGTGGTGCCCTCTCTGAGATTTTCTGAGATTTGCTCAACAG  
CCCCGATCAACATAAGTTTGTGGTACTTCACCCCTCTTGCCCTCATTTTCATGTGACAGCCATTGTGAGACTGA  
TGACAAAACCTGTCACTTGGTTAATTTAAGCACTTCTGTTTTCTGTAATTTGCTTGTTTGTTTCTTTCATGCCCTTA  
CTTACTTTGTCTCCATGCTACTGATTTGGCACGTGGGCCCCCAACATGGCACAAATAAAGCCCTTTGTGAAACTGTTC  
TTTTAATGAAACCAAGAAATTTGGCCACTGGTAAACTTCAGAGCTTCAACTGTACTTTCATTTAATGCCATTAAT  
GCAAATATATCTCCTCTTCTTTTTCATGGTTTGGCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAA  
ATAACCAATATAAAGCATATTTCTTGGCCCTTGGCTCCACAGGACATAGGCCAAGCCTTGATCATAGTTTCATACAT  
AAATCGTGTGAAATTAAGAAATAAACACATAACTTTTACTTGAATGTAAATAACTTAATTTATTTCTTTGCTA  
AATTTGGAATTTAGTGCACATTCAAAGTTAAGCTATTAATAATAGGGTGATCATAGTTCCCTCAACCAAGTCTGG  
AAAGAACATCTCCTGGTATCCACAATTTACACACAGGTTGCTAACTGTATTTGTACATTTCCCTTTGCATTCGCTTT  
TGTTCTTGTCTAGAAACCCAGTGTAGCCACGGGAGATGCTCAATAAATGCATACCTCTGTATTTTGAAAAA

**FIGURE 412**

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNID  
YPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQRP  
GQNC SNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCSEAS  
EEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTKTPKL  
LTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTPPKTSLKAATIKAEFVRAETPYMVMNP  
ETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQHQAGEYFCKA  
QSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQONATNSFYVDVGRCPVKTCAGQQ  
DNGIRCRDAVQNCCGISKTEEREIQCSGYTLPTKVAKESCQRCTETRSIVRGRVSAADNGEP  
MRFGHVYMGNSRVSMGTGYKGTFTLHVLPQDTERLVLT FVDRLQKFVN TTKVLPFNKKGS AVFHE  
IKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRQNGEPYIGKV KASVTFLDPR  
NISTATAAQTDLNFINDEGDTFPLRTYGMFSVD FRDEVTSEPLNAGKV KVHLDSTQVKMPEHI  
STVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNLDVPESRRCFVKV  
RAYRSE RFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGACVPAFCDDQSPDA  
YSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPRVKKTA FQISMAKPR  
PNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYN TVPFNEDDPMSWTE DYLAWW  
PKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRSTRDRDQPNVSAACLEF  
KCSGMLYDQDRVDRTL VKVIPQGS CRRASVNPMLHEYLVNHLPLAVNNDTSEYTMLAPLDPLG  
HNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT FNCVERQVGRQSAFQYLQ  
STPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVAQQPLIN



**FIGURE 413**

GCCACGTTGTCTTCTTTCCTTCACCACCACCCAGGAGCTCAGAGATCTAAGCTGCTTTCCATC  
TTTTCTCCCAGCCCCAGGACACTGACTCTGTACAGGATGGGGGCCGTCCTCTTGCCTCCTTCTC  
ATCCTAATCCCCCTTCTCCAGCTGATCAACCCGGGGAGTACTCAGTGTTCTTAGACTCCGTT  
ATGGATAAGAAGATCAAGGATGTTCTCAACAGTCTAGAGTACAGTCCCTCTCCTATAAGCAAG  
AAGCTCTCGTGTGCTAGTGTCAAAAGCCAAGGCAGACCGTCCTCCTGCCCTGCTGGGATGGCT  
GTCAGTGGCTGTGCTTGTGGCTATGGCTGTGGTTCGTGGGATGTTTACAGTGGAAACCACCTGC  
CACTGCCAGTGCAGTGTGGTGGACTGGACCACTGCCCCGCTGCTGCCACCTGACCTTGACAGGGA  
GGAGGCTGAGAACTCAGTTTTGTGACCATGACAGTAATGAAACCAGGGTCCCAACCAAGAAAT  
CTAACTCAAACGTCCCACTTCATTTGTTCCATTCCTGATTCTTGGGTAATAAAGACAAACTTT  
GTACCTCAAAAAAAAAAAAAAAAAAAAAA

414/550

**FIGURE 414**

MGPSSCLLLILIPLLQLINPGSTQCSLDSVMDKKIKDVLNSLEYSPSPISKKLSCASVKS  
QGRPSSCPAGMAVTGCACGYGCGSWDVQLETTCHCQCSVVDWTTARCCHLT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 415**

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGTG  
AGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAGGA  
GTGTGTGGAACAGGACCCGGGACAGAGGAACCA**ATG**GCTCCGCAGAACCTGAGCACCTTTTGCC  
TGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTGGGGG  
TGCCTCGAAGTGCCTCTATAAAGGATATTTAAAAAGGCCTATAGGAAACTAGCCCTGCAGCTTC  
ATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTGCTGCTT  
ATGAGGTTCTGTGAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAAGGATTAA  
AAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTTCTTTGGGGATTTTGGTTTCA  
TGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATATTATTGTAG  
ATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTTAGAAACAAAC  
CTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGATGCGGACCACCC  
AGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAATGCCCTAATGTCA  
AACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTGAGAGACGGCATGG  
AGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGGAGATTTACGGTTCC  
GAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATTTGTACACAAATGTGA  
CAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACTCACTTGGATGGTCACA  
AGGTACATATTTCCCGGGATAAGATCACCCAGGCCAGGAGCGAAGCTATGGAAGAAAGGGGAAG  
GGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAATCACTTTTGATGTGGATT  
TTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAACAGCTACTGAAACAAGGGT  
CAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TGA**GAGTGAATAAAATTGGACTTTGTTT  
AAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTTTTGTGTGTGTTTTTGTTTTTA  
TTTTCAATATGCAAGTTAGGCTTAATTTTTTTTATCTAATGATCATCATGAAATGAATAAGAGG  
GCTTAAGAATTTGTCCATTTGCATTTCGAAAAGAATGACCAGCAAAGGTTTACTAATACCTC  
TCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGTTTCAAGAATTAAAGCTGCAAGAGG  
ACTCCAGGAGCAAAGAAACACAATATAGAGGGTTGGAGTTGTTAGCAATTTTCATTCAAAATG  
CCAACTGGAGAAGTCTGTTTTTTAAATACATTTTGTGTATTTTTTA

**FIGURE 416**

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQAQEKFDLGAAYE  
 VLSDSEKRKQYDTYGEGLKDGHQSSHGDI FSHFFGDFGFMFGGT PRQQDRNIPRGSDI IVDLEVTLEEVYAGNF  
 VEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQMTQEVCDEC PNVKLVNEERTLEVEIEPGVRDGM EYPFI  
 GEGEPHVDGEPGDLRFRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVVHISRDKITRPGAKLW  
 KKGEGLPNFDNNNIKGLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

**Important features:****Signal peptide:**

amino acids 1-22

**Cell attachment sequence.**

amino acids 254-257

**Nt-dnaJ domain signature.**

amino acids 67-87

**Homologous region to Nt-dnaJ domain proteins.**

amino acids 26-58

**N-glycosylation site.**

amino acids 5-9, 261-265

**Tyrosine kinase phosphorylation site.**

amino acids 253-260

**N-myristoylation site.**

amino acids 18-24, 31-37, 93-99, 215-221

**Amidation site.**

amino acids 164-168

**FIGURE 417**

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGGA  
TGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGGCCGGAGAGGGCCCAGCCCCCGGGGCAG  
**GATG**ACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCTGCT  
GATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGACACGTCCTTCTCTAG  
GCCGCACACGGGGCCCGCGCTGCCCACGCCCCGGGCGGACAGGGACAGGGAGCTCACGGCCGA  
CTCCGATGTGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGACCTTCC  
CAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAGGCTACGA  
CTGGTCCCCGCGCGACGCCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAGCGGAGGAG  
CGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCGCGCATTCGA  
CGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGGCCATCTACTG  
CTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTGAGCGGAAGCCT  
GCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCACGTGCACAACGC  
CAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCTCCCGCCACCTCAT  
GAAGGTCAAGCTCAAGAAGTACACCAAGTTCTCTTCGTGCGCGACCCCTTCGTGCGCCTGAT  
CTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCGCAAGTTCGCCGTGCC  
CATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCCGCTCGGCGCGCGAGGCCTTCCGCGC  
TGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGACCCGCACACGGAGAAGCT  
GGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCACCCGTGCCAGATCGACTA  
CGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGCAGCTGCTGCAGCTACTCCA  
GGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGGACCGCCAGCAGCTGGGAGGA  
GGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCTGTATAAACTCTACGAGGCCGA  
CTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCCGAGACT**TGA**AAGCTTTCGCGTTG  
CTTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCCAGTTTTTTTTATGACCTACGATTT  
TGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCCATTGAGTACTGTATCGATATTGTT  
TTTTAAGATTAATATATTTTCAGGTATTTAATACGA

**FIGURE 418**

MTKARLFRLWLVLGSVFMILLIIIVYWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRELTAD  
SDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEEESVRGYDWSPRDARRSPDQGRQQAERRS  
VLRGFCANSSLAFTPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIIVLSGSL  
LHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPFVRLI  
SAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDPHTEKL  
APFNEHWRQVYRLCHPCQIDYDFVGKLETLEDAQAQLLQLLQVDRQLRFPPSYRNRTASSWEE  
DWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

**Important features:****Signal peptide:**

amino acids 1-31

**N-glycosylation sites.**

amino acids 134-137, 209-212, 280-283 and 370-373

**TNFR/NGFR family cysteine-rich region protein**

amino acids 329-332

**FIGURE 419**

GGCACGAGGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAAG  
GCTGCCAGGAAGGAGACGCCTTCCCTGAGTCCTGGATCTTTCTTCCCTTCTGGAAATCTTTGACT  
GTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATGG**ACCTCGCGGGACTGCTGA  
AGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGGCTAATCATCA  
ACACCATTTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAAGATCA  
ACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGTCGGGCA  
CGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAATGCCATCG  
TGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGAACGCTTTG  
GGCTGTTAGGGGGCTCCAAGGTCCTGGCCAAGAAAGAGCTGGCCTATGTCCCAATTATCGGCT  
GGATGTGGTACTTCACCGAGATGGTCTTCTGTTTCGCGCAAGTGGGAGCAGGATCGCAAGACGG  
TTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCTGATTCACTGTG  
AGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCCGGGCCAAGGGGC  
TGCTTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATCACCGTGAGGAGCT  
TGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAATAATGAAAATCCAA  
CACTGCTGGGAGTCTTAAACGGAAAGAAATACCATGCAGATTTGTATGTTAGGAGGATCCAC  
TGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCACAAGCTCTACCAGGAGA  
AGGATGCCTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGAGACGCCCATGGTGCCCC  
CCCGGCGGCCCTGGACCCTCGTGAACCTGGCTGTTTTGGGCCTCGCTGGTGCTCTACCCTTTCT  
TCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACGCTGGCCAGCTTCATCCTCG  
TCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGTGACGGAAATTGACAAGGGCT  
CTGCCTACGGCAACTCTGACAGCAAGCAGAAACTGAATGACT**TGA**CTCAGGGAGGTGTCACCAT  
CCGAAGGGAACCTTGGGGAACCTGGTGGCCTCTGCATATCCTCCTTAGTGGGACACGGTGACAA  
AGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGACCTCTCCAGCCAGGGAGTCTGGT  
CTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTTTTTCCCCATGTGCTTTAGTGGGC  
TTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGCTGTGTGGTGAGTGTGAACTTTGTTT  
TGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAGGGCAGGGCTGGGGACCGAAGGGGACA  
AGTTCCCTTTTCATCCTTTGGTGCTGAGTTTTCTGTAACCCTTGGTTGCCAGAGATAAAGTGA  
AAAGTGCTTTAGGTGAGATGACTAAATTATGCCTCCAAGAAAAAAAAAATTAAAGTGCTTTTCT  
GGGTCAAAAAAAAAAAAA

**FIGURE 420**

MDLAGLLKSQFLCHLVFCYVFIFASGLIINTIQLFTHLLWPINKQLFRKINCRLSYCISSQLVM  
LLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHNKFEIDFLCGWSLSERFGLLGSKVLAKKEL  
AYVPIIGMMWYFTEMVFCSRKWEQDRKTVATSLQHLRDYPEKYFFLIHCEGTRFTEKKHEISM  
QVARAKGLPRLKHLLPRTKGFATVRSRLNRNVSAYVDCTLNFRNNENPTLLGVLNGKKYHAD  
LYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVNWLFWA  
SLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDSKQKLN



**FIGURE 421**

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCATC  
GCC**ATG**GACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGCCC  
TGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTCACC  
ACAGTCCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGCGGGC  
CTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGGCGCTGGGTGCC  
CTGAAGGAGGAGGTTCGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTGCAGACC  
ACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCTGCGGGAA  
CTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACGTCCGCACT  
GAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACCTCCTGCGAGCCGTGCCCCACG  
TCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTTCTCTGTGCCAAAGACGACGTGGGCGGGC  
GCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCCCTGGATGAGCAG  
GGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGGGCTGTGCGCCAT  
CTGGGCAAGGTTTCAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTTCAGCCACTGGAAC  
CAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTGTCATGATGCTGCACACGGGGCTG  
TGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAGAAAAGGCACAACCTGC  
**TGA**CCCCGCCAGTGCCCTGGAGCCGCGCCCATTCGAGCATGTCGTATCCTGGGGGCTGCTCA  
CCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTTCTTCCTCATCCACCGCTGCTGAG  
TCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCCTGGGCTCTGGGACCTCCA  
TGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACCTCCACTAGCTCCAAAATCC  
CTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAACCAAGGTTAGGTGACTGAGG  
ACTGGAGCTGTTTTGGTTTTCTCGCATTTTCCACCAAACCTGGAAGCTGTTTTTGCAGCCTGAGG  
AAGCATCAATAAATATTTGAGAAATGAAAAAA

**FIGURE 422**

MDTTRYSKWGGSSSEVPGGPWGRWVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAAL  
LDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALREL  
RERV TQGLAEAGRGREDVRTELFRAL EAVRLQNN SCEPCPTSWLSFEGSCYFFSVPKTTWAAA  
QDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSF SHWNQ  
GEPNDAWGREN CVMLHTGLWNDAPCDSEKDGWICEKRHNC

**Important features:****Type II transmembrane domain:**

amino acids 31-54

**N-glycosylation sites.**

amino acids 73-76 and 159-162

**Leucine zipper pattern.**

amino acids 102-123

**N-myristoylation sites.**

amino acids 18-23, 133-138 and 242-247

**C-type lectin domain signature.**

amino acids 264-287

**FIGURE 423**

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGGC  
 GAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCGGC  
 GCCCAAC**ATG**GCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGATCGC  
 GGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGGTCCA  
 GCCCATGACCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTTTACGC  
 CCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGCTTTTGCAAAGAATGGTGA  
 AATACTTCAGATCAGTGTGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTGGCCGCTT  
 CTTTGTCACTCTCTCCAGCATTTTTTTCATGCAAAGGATGGGATATTCCGCCGTTATCGTGG  
 CCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATCAGTCGAGCC  
 TCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTCTTTTTAGCAT  
 CTCTGGCAAGATATGGCATCTTCACAACATTTTACAGTGAAGTCTTGGAAATTCCTGCTTGGTG  
 TTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCTTTTTATGGGTCTGGTCTTGGT  
 GGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGCGTTCTGAGCAGAA  
 TCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAGGAGGAAAAAGATGA  
 TTCAAATGAAGAAGAAAAACAAAGACAGCCTTGATAGATGATGAAGAAGAGAAAGAAGATCTTGG  
 CGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTTGGCTGCTGGTGTGGATGAGGA  
 GAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTGACCCGGGAGGAAGTAGA  
 GCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCAGCTGACACAGAGGTGGTGGGA  
 AGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGACTG**TAG**ATTTAATGATGCGT  
 TTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGGCCTGCAGTTTGTACCAAATC  
 CTTAATTTTTCTGAATGAGCAAGCTTCTCTTAAAAGATGCTCTCTAGTCATTTGGTCTCATG  
 GCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGTGACAATCAGGATATAGAAAAAC  
 AAACGTAGTGTTGGGATCTGTTTGGAGACTGGGATGGGAACAAGTTCATTTACTTAGGGGTCA  
 GAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCTTAATCAGCACCTTCCAGAGACAAGGCTGC  
 AGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCTCCTGAGCATCCCCAAAGTGTAACGT  
 AGAAGCCTTGTCATCCTTTTCTTGTGTAAAGTATTTATTTTTGTCAAATTGCAGGAAACATCAG  
 GCACCACAGTGCATGAAAAATCTTTCACAGCTAGAAATTGAAAGGGCCTTGGGTATAGAGAGC  
 AGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTGTGCTATGTTTTATTTCTTACCTTTAATT  
 TTTCCAGCATTTCCACCATGGGCATTACAGGCTCTCCACACTCTTCACTATTATCTCTTGGTCA  
 GAGGACTCCAATAACAGCCAGGTTTACATGAAGTGTGTTTGTTCATTCTGACCTAAGGGGTTT  
 AGATAATCAGTAACCATAACCCCTGAAGCTGTGACTGCCAAACATCTCAAATGAAATGTTGTG  
 GCCATCAGAGACTCAAAGGAAGTAAGGATTTTACAAGACAGATTAAAAAAAATTGTTTTGT  
 CCAAATATAGTTGTTGTTGATTTTTTTTTTAAGTTTTCTAAGCAATATTTTTCAAGCCAGAAG  
 TCCTCTAAGTCTTGCCAGTACAAGGTAGTCTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCT  
 ATCTCAAGGGGTTCCCTGGGTCTTGAAGTACTTTAATAATAACTAAAAAACCACTTCTGATTT  
 TCCTTCAGTGATGTGCTTTTGGTGAAAGAATTAATGAAGTCCAGTACCTGAAAGTGAAAGATT  
 TGATTTTGTTCATCTTCTGTAATCTTCCAAGAATTATATCTTTGTAAATCTCTCAATACT  
 CAATCTACTGTAAGTACCCAGGGAGGCTAATTTCTTT

## **FIGURE 424**

MAGGRCGPQLTALLAAWIAAVAATAGPEEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAPW  
CPSCQQTDSEWEAFKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGI FRRYRGPG  
IFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAWCSY  
VFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSESEQNRRSEEAHRAEQLODAEEKDDSN  
EEENKDSLVDDEEEKEDLGDEDEAEEDNLAAGVDEERSEANDQGPPGEDGVTREEVEPE  
EAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **Transmembrane domain:**

amino acids 191-211

#### **N-glycosylation site.**

amino acids 46-49

**Thioredoxin family proteins.** (homologous region to disulfide isomerase)

amino acids 56-72

#### **Flavodoxin proteins**

amino acids 173-187

[illegible]

**FIGURE 426**

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTQDFYSLLGVSKTASSREIRQAFKKLALKLH  
 PDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYRYDFGIYD  
 DDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNCGDD  
 RMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNSIQTA  
 FAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSANTLEDR  
 LAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQPSLAVFK  
 GQGTKEYEIIHHGKKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWCPPCRALL  
 PELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEYEGHHSAEQILEFI  
 EDLMNPSSVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMARTLTGLINVGS  
 IDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLGFLPQVSTDLT  
 PQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVKAGKVCQAYAQTCQKAG  
 IRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAALISEKLETLRNQGKRNKDEL

**Important features:****Endoplasmic reticulum targeting sequence.**

amino acids 744-747

**Cytochrome c family heme-binding site signature.**

amino acids 158-163

**Nt-dnaJ domain signature.**

amino acids 77-96

**N-glycosylation site.**

amino acids 484-487

**FIGURE 427**

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGCA  
CGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTTGA  
GTCCTTCTGAG**ATG**ATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGCATGGTAG  
CGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTCAATT  
CCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTCTGCAG  
TCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACAACTACC  
AGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGTCCCACCC  
GCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACGCTGCATGC  
GTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTTCTGATCAA  
ATCATTTCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATAGCACCT  
TGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAAAGGACAAGAAG  
GTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTAGACACTTCTGGT  
CCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCATAGGAGAAAAGGCT  
CTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTCTTGCCGGATACAGA  
AAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGAGACAC**TAA**ACCAGCT  
ATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACCTTTTATGACCTTCATC  
AACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCATTCCAATAACACCTTCCA  
AAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCTGTGATTGCAGTAAATTACT  
GTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGAACTTTTAATTATTTTCT  
AAAGGTGCTGCACTGCCTATTTTTCCTCTTGTTATGTAAATTTTTGTACACATTGATTGTTAT  
CTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATTTTCAGCTTATAGTTCTTAAAAG  
CATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCAAGGATCTCTTGGAATGACAAAT  
GATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATTTTCTGAAATGTACTATCTTAATG  
CTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGAAATAAAATTTAACATTTAAAAAA  
AAAAAA

428/550

## **FIGURE 428**

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSAA  
PGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRCMRHAM  
CCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTLSSKMYHTKGQEGSVC  
LRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQKDDH  
QASNSSRLHTCQRH

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **N-glycosylation site.**

amino acids 256-259

#### **Fungal Zn(2)-Cys(6) binuclear cluster domain**

amino acids 110-126



**FIGURE 429**

GAGAGGACGAGGTGCCCTGGAGAATCCTCCGCTGCCGTCCGGCTCCCGGAGCCCAGCCC  
TTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGACCC  
CAGCGTTACCA**ATG**CATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCT  
CCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGAATAT  
AGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGTCGTTT  
CAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATTTCCAAA  
TGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATA  
CAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATA  
CAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAGTGACCCCAT  
TCAAGAAATTCGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAAGAAATATCATTGG  
ATATTTTGAGCAAAAGGACTCGGACAACCTATAGAGTTTTTGAACGAGTAGCGAATATTTTGCA  
TGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAACCCGAAAGATATAGTGGCGA  
CAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACTTGGGAGCTATGAC  
AAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCCTCTTGTCCGAGAAATAAC  
ATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCATACTCTTTCACATGAA  
AGAAGATACAGAAAGTTTAGAAATATTCAGAATGAAGTAGCTCGGCAATTAATAAGTGAAAA  
AGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACATCCTCTTCTGCACATACA  
GAAAACCTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAGGCATATGTATGTGTTTGG  
AGACTTCAAAGATGTATTAATTCCTGGAAAACCTCAAGCAATTCGTATTTGACTTACATTCTGG  
AAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACTGATACAGCCCCAGGAGAGCAAGC  
CCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAAACTAGCACCCAGTGAATATAG  
GTATACTCTATTGAGGGATCGAGATGAGCTTT**TAAA**AACTTGAAAAACAGTTTGTAAGCCTTTC  
AACAGCAGCATCAACCTACGTGGTGGAATAGTAAACCTATATTTTCATAATTCTATGTGTAT  
TTTTATTTTGAATAAACAGAAAGAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAA

## **FIGURE 430**

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQM  
LHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYRGQ  
RSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILHDDC  
AFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREITFEN  
GEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLHIQKTP  
ADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPGEQAQDV  
ASSPPESSFQKLAPSEYRYTLLRDRDEL

**Important features:**

**Signal peptide:**

amino acids 1-29

**Endoplasmic reticulum targeting sequence.**

amino acids 403-406

**Tyrosine kinase phosphorylation site.**

amino acids 203-211

**Thioredoxin family proteins**

amino acids 50-66

**FIGURE 431**

GAGCAGGACGGAGCCATGGACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGCA  
GGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCGTG  
CAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGCGTG  
GACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGCAGTG  
CGGGGTTCGCGTTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGCTTCTG  
GCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTCACCTCG  
CGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTGCTACAGC  
TGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCTGCTACAAC  
GCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCAGCTAATGTG  
ACTGTGTCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGATGGAGTAACA  
GGCCCAGGGTTACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACCTTGACCTCCGC  
AACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCTCCAGAGCCCACG  
ACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAGACCCACATCCACC  
ACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAGAACACGAGGCCTCC  
CGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGACCGCAGCAATTCAGGG  
CAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTGTGTGGCTCCCACAGCT  
GGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTTACTGTGAGGCTTCTCCACCTGGA  
AATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCTCATCACTTCCTGTTCCCA  
CCACTGGACTGGGCTGGCCAGCCCCTGTTTTTCCAACATTCCCCAGTATCCCCAGCTTCTGC  
TGCGCTGGTTTTGCGGCTTTGGGAAATAAAATACCGTTGTATATATTCTGCCAGGGGTGTTCTA  
GCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTCTCCGCTTGTCTCTTGTGATG  
TTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTGAGAGAGAGGATGCTAAGCTTCC  
TACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGGGGTGGGTGGGACAATGGCTCCCC  
ACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGGGAATCGGTTCCCCATATGTCTTCC  
TTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTACCCAATTCGCCCTATAGTGAGTCGTA

432/550

## **FIGURE 432**

MDPARKAGAQAMIWTAGWLLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCTE  
AVGAVETIHGQFSLAVRGCGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRALDP  
AGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCDFGNVTTLTAANVTVSLP  
VRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTTVAST  
TSVTTSTSAFVRPTSTTKPMPAPTSQTPRQGEHEASRDEEPRLTGGAAGHQDRSNSGQYPAK  
GGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

432/550

FIGURE 433

[illegible]

434/550

**FIGURE 434**

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRRCKC  
SFNQKPRAPGDDEEAQVENLITANATEPQKQRTevQPSGGSLWNLRRLLLEPLDANVDA

434/550

**FIGURE 435**

GGTCCTTA**ATG**GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTGC  
TGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGTCA  
TCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGACTT  
TTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAACTAA  
ATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACTTACAG  
AGCAACTGCGTGACATTGAGCTGGAGAATTACACACCCAAGGAACCCCTCACCTGCAGGCAA  
GGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGTTTCGATG  
GGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCCTGGAGCCA  
GAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATTACTTCTCAA  
TGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACCCCTGGAGCCAA  
GTGCAGGAGCACCCTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCACAGCCACCACCC  
TCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCATCTGAGGAGAGT  
CCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCACGGTCTTGATCAAAC  
TCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGTGGCCTCCAGCAGAT  
CATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTTTTGCCAACAATTTTA  
CCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACCTGATGGAATTCCTGCA  
CTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTCTTTTTGTTTGGAAAA  
TCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATTGTCAGTAAAATAATCACG  
TTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAGAATTTTAAATTATTTAAT  
AAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTATTGTTCTGTACTGATATTTAA  
ATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

436/550

## **FIGURE 436**

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVI PKFRPGPRWCAVQGQVDEKTF  
YDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQARMS  
CEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFSMGD  
CIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 437**

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACCATGAAGCTCTTGTGGCAGG  
 TAACTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTCTGTCTACCTCACGGCGCAAGTGTGGATTCTGT  
 GTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCCAGAACTGCCCCCTCCGTTTGTCTGTGCAGTAACCAGTTCA  
 GCAAGGTGGTGTGCACGCGCCGGGGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACC  
 TCATGGAGAACAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGCAGTTGG  
 GCAGGAACCTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAACACCCTGGAGCTGTTTCG  
 ACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTCCAAGCTGCGGGAGCTCTGGCTTCGCAACA  
 ACCCATCGAAAGCATCCCCTCTTACGCCTTCAACCGGGTGCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCA  
 AGAAGCTGGAGTATATCTCTGAGGGAGCTTTTGGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCA  
 ACATTAAAGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTGAGGGAACCACTTCCCTG  
 AGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCTGAACTCACAGGTGAGCCTGA  
 TTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAACCAACTTGGCCCCACAATAACCTCTCTTCTTTGC  
 CCCATGACCTCTTTACCCCGCTGAGGTACCTGGTGGAGTTGCATCTACACCACAACCCCTTGGAACTGTGATTGTG  
 ACATTCTGTGGCTAGCCTGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTGATGCTC  
 CCATGCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCTTCATCATGG  
 ACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGTGCGACTCCCCCTATGTCTCCG  
 TGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGCCTCCCGCCACCCAAGGATCTCTGTCTCAACGACG  
 GCACCTTGAACTTTTCCACGTGCTGCTTTAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCA  
 ACTCCAACGCCTCGGCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTACCACAG  
 TAACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTTCTACCACGTCCACTG  
 GTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTGAGACTACCCGTGTGCCCAAGCAGGTGGCAGTAC  
 CCGCGACAGACACCACTGACAAGATGCAGACCAGCCTGGATGAAGTCATGAAGACCACCAAGATCATCATTGGCT  
 GCTTTGTGGCAGTGAATCTGCTAGCTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGC  
 GGAGTACAGTCACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGCAG  
 CAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCAACAATTGATGACCATATTAAC  
 ACAACACCTACAAACCAGCACATGGGGCCCCACTGGACAGAAAACAGCCTGGGGAACTCTCTGCACCCCCACAGTCA  
 CCACTATCTCTGAACCTTATATAATTGAGACCCATACCAAGGACAAGGTACAGGAACTCAAATATGACTCCCCCT  
 CCCCCAAAAAATTATAAAATGCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTT  
 TTCTTGTATATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAATAATTTAAAGACAAAA  
 AGTCAAAACA

**FIGURE 438**

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCTR  
RGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLASLN  
TLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEYISE  
GAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEEELEMSGNHFPFIRPGSFHGLSSLKKLWVMNS  
QVSLIERNAFDGLÅSLVELNLAHNNLSSLP HDLFTPLRYLVELHLHHNPWNCDCDILWLAWWL  
REYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAELKCRTPP  
MSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTG VYTCMVTNVAGNSNASAYLNV  
STAE LNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQTTRVPKQV  
AVPATD TTDKMQTS LDEV MKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRSTVTAARTVE  
IIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHW TENS LGNSLHPT  
VTTISEPYIIQTHTKDKVQETQI

**FIGURE 439**

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGAGGCTG  
TCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCTGC  
CCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTCCAA  
GTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTGCACC  
GATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAATTAGTGAAAAAAT  
GTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACCCTGATC  
TTCATAAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

440/550

**FIGURE 440**

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNQLQVAKLNPPPEALAAKLEVK  
HCTDQISFKKRLSLKKSWWK

440/550

**FIGURE 441**

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGATG  
GGGTTGCTGGTTTAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCACC  
ACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCC**ATG**CCCTCCCCAGG  
GACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCTCCAG  
CTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCAGCCAA  
GCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGAAGGGGC  
AGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGTCAGGGGT  
TCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGGGAAGAGGC  
CAAAGAGGCCCCAGCCGACAAG**TGA**TCGCCCACAAGCCTTACTCACCTCTCTCTAAGTTTAGA  
AGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTACAAGCTCAGG  
AGGCGAATAAATGTTCAAACCTGTA

442/550

**FIGURE 442**

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDGG  
QAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

442/550

**FIGURE 443**

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTCG  
TGAACCCCGGGGTGCTCCGCACGGACCCAGATGTCAAGAATATGAACACGTGGCTGCTGTTC  
CTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTGCTC  
CTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCTGAGT  
ATGTCCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATCTTACA  
CCTCTACTTGAGTATGTCCCTAACCTGAGCCCCCACGCCTGGGGCCAGAGTCTTTGTCCCC  
CGTGTGCGCATGTGTTTCAGGGTCAGCCTCTCCAGAAGTGAGATCATGGACAAAAAGGGCAA  
TCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCAGCAAGAAGCTGAACTCACGCCG  
AGACCTGCAGGAGTGGTGCCAGGTGCTTGAAGTAACAAGTTTAAATGTTTCAGAGACAATGGA  
ATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTTCCAAAAACAC  
AAGTAGAAATTCTAACAAATGAAATATATTACAGGCAGGTCACCCACTAACCAACAACCTGAAG  
CGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTTCAGTCATGTTGCT  
GAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGTAACAACAACCTCC  
CTGCTCCTGGCACCAGCCGTTTTGGTTCATGGTGGGCCAGCTGCAAAGCGTCTTCCATTCTCTG  
GGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGACACGGGCAGCAGAGTG  
TGTCCAGGCCAGCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCAACCCCTGGTCAGGGCA  
GAGGGAGTTGGGTGGGTTCAGGCTCTGGGCTCACCTCCATCTCCAGAGCATCCCCTGCCTGCAG  
TTGTGGCAAGAACGCCCAGCTCAGAATGAACACACCCCCACCAAGAGCCTCCTTGTTTCATAACC  
ACAGGTTACCCTACAAACCACTGTCCCCACACAACCCTGGGGATGTTTTAAACACACACCTC  
TAACGCATATCTTACAGTCACTGTTGTCTTGCTGAGGGTTGAATTTTTTTTAAATGAAAGTGC  
AATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAA

444/550

**FIGURE 444**

MNTWLLFLPLFPVQVQTLIVVIIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQGWV  
VRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQAQ  
QEAELTPRPAGVVPGA

444-444



**FIGURE 445**

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGA**ATG**GACTGGCCTCACAACTGC  
TGTTTCTTCTTACCATTTCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAAGA  
GGAAGGGGCAAGGGCGGCCTGGGCCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACCTGG  
TGTCACGGATGAAACCGTATGCCCCGATGGAGGAGTATGAGAGGAACATCGAGGAGATGGTGG  
CCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCTGTGGA  
TGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCCGTATCC  
CCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACCATGCAGG  
AGGACCGCAGCATGGTGAGCGTGCCGGTGTTTCAGCCAGGTTCTGTGCGCCGCCGCCTCTGCC  
CGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCGCTGTGGGCT  
GCACCTGCATCTTCT**TGA**ATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGACCATCCTCCT  
TGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAG

## **FIGURE 446**

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEYE  
RNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPPWGYSINHDPSRIPVDLPEARCLCLGC  
VNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **N-glycosylation site.**

amino acids 75-78

#### **Homologous region to IL-17**

amino acids 96-180.

446/550

**FIGURE 447**

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC**ATG**GCCAAGA  
TGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCTAT  
CACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGGTGC  
CCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCTGGATG  
GAGATACCAACACATCCACCCAGGAGGTGGTACAATAACAAGTGGGAGACTGGGGATGACCGGT  
TCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAGAACCAG  
GGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAAGGACTAC  
TGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAAGCGGTTGA  
TGGAGAAGGCTTCCCTCCCCTCCCCTCCCCTTGGGGCTTTGTGGCAAAAATCCTATGGTTATCC  
CTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCTACTAACAGAC  
TTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAAGTGGAGCGCCTTTGCTGCTGTTTCCTCT  
GTCCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTACAAAGTCTTCCAAGCGACT  
GTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTGGGCCTTCTACATG  
GCCTGGCTCTCCTTACCTGCTGCATGGCGTCGGCTGTCACCACCTTCAACACGTACACCAGG  
ATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAACCCGAAGTGCCTACCA  
CATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCACCGTGGGTCCCTTTGACC  
AGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGGGAGTCGACTTCTACTCC  
GAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTGAAAGAAGCAGTTAGGTCA  
TCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGAGTAGGCTTGAGCCCTACCT  
TACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCGTCTCTTGAGCATGGTTTTTA  
GAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTCCTAAGGGATTCTGGGTGCCA  
CTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTACCCACCCACATCTCACACATCCAGAA  
TTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGCTAAACCATGGAGATAAAAAGAAG  
AGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

**FIGURE 448**

MAKMELSKAFSGQRTLLSAILSMLSLSFSSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMPV  
SLDGDNTNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKRGE  
KGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPLGLCGKNPMVIPGNADHLHRTSIHQLPPA  
TNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSCQLGSRRLTTCLELWLG  
LLHGLALLHLLHGVGCHHLQHVVHQDGAGVQVQA

448/550

**FIGURE 449**

CCCACGCGTCCGCCCCACGCGTCCGCCCCACGCGTCCGCCCCACGCGTCCGCCCCACGCGTCCGCCCC  
 ACGCGTCCGCGTGCAGCTCGCGCCGCACACTGCCTGGTGGAGGGAAGGAGCCCGGGCGCCTCTCGCCGCTCCCGG  
 CGCCGCGTCCGCACTCCCCACGCGCCCGCCCGCCCGCCCGCCGCAAGCATGAGTGAGCCCGCTCTCT  
 GCAGCTGCCCCGGGGCGCAATGGCAGGCTGTTTCCGCGGAGTAAAGGTGGCGCCGGTCAGTGGTCTTTCCAAT  
 GACGGACATTAACCAGACTGTAGATCCTGGGGAGTCGCGAGCCCCGAGTTTGGAGTTTTTCCCCCACAACGT  
 CACAGTCCGAACGTCAGAGGGAAAGGAAGGCGGCAGGAAGGCGAAGCTCGGGCTCCGGCACGTAGTTGGGAAACT  
 TGGGGTCTAGAAAGTCGCTCCCCGCTTGGCGGGCCGCTTGCAGCCCCGAGCCGAGCAGCAAGTGAGACAT  
 TGTGCGCTGCCAGATCCGCGCGCCGCGGACCGGGCTGCCTCGGAAACACAGAGGGGTCTTCTCTCGCCCTGCA  
 TATAATTAGCCTGCACACAAAGGGAGCAGCTGAATGGAGGTTGTCACTCTCTGGAAAAGGATTCTGACCGAGCG  
 CTTCCAATGGACATTCTCCAGTCTCTCTGGAAAGATTCTCGCTAATGGATTCTCTGCTGCTCGGTCTCTGTCTAT  
 ACTGGCTGCTGAGGAGGCCCTCGGGGTGCTTGTGTCTGCTGGGGGCTGCTTTCAGATGCTGCCCCCGCGCC  
 CCAGCGGGTGCCCGCAGCTGTGCCGTTGCCAGGGGCGGCTGTGTACTGCGAGGCGCTCAACCTCACCAGGCGC  
 CCCACAACCTGTCCGGCTGCTGGGCTTGTCCCTGCGCTACAACAGCCTCTCGGAGCTGCGCGCCGGCCAGTTCA  
 CGGGGTTAATGCAGCTCAGTGGCTCTATCTGATCACAATCAGATCTGCTCCGTGCAGGGGGACGCTTTTCA  
 AACTGCGCCGAGTTAAGGAACCTCAGCTGAGTTTCCAACAGATCACCCTGCTGCGGCGCAACTCTTTCGCGGCTTGT  
 TGCCCAACCTGCGCAGCGTGGACCTCTCTGTAACAAGCTGCAGGCGCTGCGCGCCGACCTCTTCCACGGGCTGC  
 GGAAGCTCACCAGCTGCATATCGGGGCAACGCCATCCAGTTTGTGCCGTTGCGCATCTTCCAGGACTGCCGCA  
 GCCTCAAGTTTCTCGACATCGGATACAATCAGCTCAAGAGTCTGGCGCGCAACTCTTTCGCGGCTTGT  
 TCACCGAGCTGCACCTCGAGCACAACGACTTGGTCAAGGTGAAGTTTCCGCCACTTCCGCGCCTCATCTCCCTGC  
 ACTGCTCTGCTGCGGAGGAACAAGGTGGCCATTGTGGTCAAGTCTGCTGGACTGGGTTTGGAACTGGAGAAAA  
 TGGACTTGTGCGGCAACGAGATCGAGTACATGGAGGCGGCTGTGTACTGAGACCGTCCGCGACCTGCAGTCCCTGC  
 AGCTGGACTCCAACCGCTCAGCTACATCGAGGCGGCTGCTGCTGAGAGCTGCGCGCGCAACTCTTTCGCGGCTTGT  
 TGCCCGGGAACCTGTGGGATTGCGGGCGCAACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
 ACCGATGGCAACTTGCAGTGCAGCAGCCGAGTACGACAGGGGCGAGGACGTCCTGGACGCGGTGACGCTTCC  
 ACCTGTGCGAGGATGGGGCCGAGCCACAGCGGCCACCTGCTCTCGGCCGTCAACACCGCAGTGATCTGGGGC  
 CCGCTGCCAGCTCGGCCACACGCTCGCGGACGCGGGGAGGGGAGGAGCAGCAGCGGCACATTGAGCCTGCCACCG  
 TGGCTCTTCCAGGCGGCGAGCAGCGCGAGAAGCGCGTGCAGTCCACAAGGTGGTACGGGACCATGGCCCTCA  
 TCTTCTCTTCT  
 AGGAATACTACGTTGATTACAAACCGAACACATTGAGGGAGCCCTGGTGTATCATCAACGAGTATGGCTCGTGT  
 CCTGCCACAGCAGCCCGCGAGGGAATGCGAGGTGATGTTGCTCCAGTGGCTCTCAACCCATGCGCTACCAAATA  
 CGCCTGGGCGAGCCGGGACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG  
 TGAAACTTTAAGGGGATCTCTCCAGAGACTTGACATTTTAGCTTTTATTGTTGCTTAAAAACAAAAGCGAATTAA  
 AACACAACAAAAACCCACCCACACCTTCCAGGACAGTCTATCTTAAATTTTATATGAGAATCCTTCTCTCC  
 TTTGAAGATCTGTCCATATTAGGAATCTGAGAGTGTAAAAAAGGTGGCCATAAGACAGAGAGAGAATAATCGTG  
 CTTTGTTTTATGCTACTCTCTCCACCTGCCCATGATTAAACATCATGTATGTAGAAGATCTTAAAGTCCATACGC  
 ATTTTCATGAAGAACCATTGGAAGAGGAATCTGCAATCTGGGAGCTTAAGAGCAAAATGATGACCATAGAAAGCTA  
 TGTTCTTACTTTTGT  
 AACGGGAATTTAGCTCACATCATTTTCATGCCCCTGTGCTCTAGCTCTGGAGATTGGTGGGGGGAGGTGGGGGGA  
 AACGGCAGGAATAAGGGAAAGTGGTAGTTTTAACTAAGGTTTTGTAACACTTGAAATCTTTTCTTCTCAAATTA  
 ATTATCTTTAAGCTTCAAGAACTTGTCTGACCCCTCTAAGCAAACTACTAAGCATTAAAAAGAGAATCTAATT  
 TTTAAGGTGTAGCACCTTTTTTTTTTATTCTTCCACAGAGGGTGCTAATCTCATTATGCTGTGCTATCTGAAAA  
 GAACTTAAGGCCACAATTCAGTCTCGTCTCGGCGATTGTGATGGATTGACCCCTCCATTTGACGTACCTTCCAG  
 CTGATTAAAGTTTCAAGAGTGGTATTGAGGTTTTTCAATATTTATATAGAAAAAAGTCTTTTTCATGACAAAT  
 GACACTCTCACACAGTCTTAGCCCTAGTAGTTTTTTAGGTTGGACCAGAGGAAGCAGGTTAAATGAGACCTGTC  
 CTCTGCTGCACTCAGAAAAAATAGGCACTCCCTGATGCTCAGATCTTAGCCTTGATATTAATAGTTGAGACCACC  
 TACCCACAATGCAGCCTATACTCCCAAGACTACAAAGTTACCATCGCAAGGAAAGGTTATTCCAGTAAAGGAA  
 ATAGTTTTCTCAACATTTAAAAATATTCTTCTGAACCTCATCAAAGTAGAAGAGCCCCCAACCTTTTCTCTCTGC  
 CTTCAAGAAGGCAGACATTTGGTATGATTTAGCATCAACAACACATTTATGAGTATATGTAAGTAATCAGAGGGG  
 CAAATGCCACTTGTATTCTCCCAAGTTTTTCAAGCAAGTACACACAGATCTCTGGTAGGATTAGGGGCCACTT  
 GTGTTTCCGGCTTATTTTAGTCGACTTGTACGAAGTTTGTGCTAGTCTATCTGACATGGCCAGTAGAACAG  
 GGCATTGATGGATCACATGAGATGGTAGAAGGAACATCATCACATACCCCTCTCACAGAGAAAATTATCAAGAA  
 CCAGAAATTTATCTGTTTTGGAGCAAGAGTGTATAATGTTTCAGGGTAGTCAAAATAAACATAAATTATCTCC  
 TCTAGATGAGTGGCGATGTTGGCTGATTTGGGTCTGCCATTGACAGAATGTCAATAAAAAAGGAATTAGCTAGAA  
 TATGACCATTAAATGTGCTTCTGAAATATATTTTGGATAGGTTTAGAATGTCA

**FIGURE 450**

MDFLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPHN  
LSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQIT  
QLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIQDCRSLKF  
LDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLISLHSLCLRRNKVAIVVSSL  
DWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKS LTSITLAGNL  
WDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGAEP TSGHLLSAV  
TNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMALIFSFLI  
VVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNHIEGALVII  
NEYGSCTCHQQPARECEV

**FIGURE 451**

TTGAGCGCGCAGGTGAGCTCTCTGCGCGTTCCGGGGGGCGTTCCCTCCAGTCACCCCTCCCGCCGTTAC  
CCGCGGCGCGCCCCGAGGGAGTCTCTCCAGACCCTCCCTCCCGTTGCTCCAAACTAATACGGA  
CTGAACGGATCGCTGCGAGGGTGGGAGAGAAAATTAGGGGGAGAAAGGACAGAGAGAGCAACT  
ACCATCCATAGCCAGATAGATTATCTTACACTGAACTGATCAAGTACTTTGAAAATGACTTTCG  
AAATTTTATCTTGGTGTCTTCATACTTTGCTGCACTGAGTCTTTCAACCACCTTTTCTCTCCAA  
CTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACTTATATAAA  
GTTCCAACGCCCCATTTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAAT  
GTTTTTATTACAAAACCTACCCTAACCATTATACTTTGGTAACCTGGCCTCTTTGCAGAGAAT  
CATGGGATTGTTGCAAATGATATGTTTGATCCTATTCGGAACAAATCTTTCTCCTTGGATCAC  
ATGAATATTTATGATTCCAAGTTTTGGGAAGAAGCGACACCAATATGGATCACAAACCAGAGG  
GCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGAACAGATGTAAAAATACATAAGCGCTTT  
CCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAGAGTTGCCAAAATTGTT  
GAATGGTTTTACGTCAAAGAGCCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTGATGAC  
ATGGGGCCACCATTTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCAATTCAGATATTGACAAG  
AAGTTAGGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATC  
ATCACAAGTGATCATGGAATGACGCAGTGCTCTGAGGAAAGGTTAATAGAAGTTGACCAGTAC  
CTGGATAAAGACCACTATAACCCTGATTGATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAA  
GGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCTCATCCTAATCTTACTGTTTACAAA  
AAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTC AACCAATCATAGCA  
GTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACCAC  
GGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTTAGCCCATGGTCTGCTTCAGAA  
AAGAATTTCTCAAAAAGAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTC  
AATATCACTGCCATGCCACACAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCA  
ATGCCAAGGGTGGTCCCTTATACACAGAGTACTATACTCCTCCCTGGTAGTGTTAAACCAGCA  
GAATATGACCAAGAGGGGTGCATACCCTTATTTCATAGGGGTCTCTCTTGGCAGCATTATAGTG  
ATTGTATTTTTTTGTAATTTTTCATTAAGCATTTAATTACAGTCAAATACCTGCCTTACAAGAT  
ATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAATGTTACTTTGAAGTGGATTTGCAT  
ATTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATCTGGGAAACCAGTT  
CCAAACATCTGCAGAAACCATTAAGCAGTTACATATTTAGGTATACACACACACACACACACA  
CACATACACACACACGGACCAAATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGT  
CTCCATTGTTCACTGTAGCATAGGGATAGATAAGATCCTGCTTTATTTGGACTTGGCGCAGAT  
AATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATATATTGCACTTTAAATTTCT  
CTCCTGATGGGTACTTTAATTTGAAATGCACCTTTATGGACAGTTATGTCTTATAACTTGATTG  
AAAATGACAACCTTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACCTGAAGGA  
AATTTCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAA  
GGTGATAAGTGTTGAAAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCC  
CAACAGCAGAATGCAACTGTGGGCATTTCTTGTCTTATTTCTTTCCAGAGAACGTGGTTTTTCA  
TTTATTTTTTCCCTCAAAAGAGAGTCAAATACTGACAGATTCGTTCTAAATATATTGTTTCTGT  
CATAAAATTTATTGTGATTTCTTGATGAGTCATATTACTGTGATTTTCATAATAATGAAGACAC  
CATGAATATACTTTTCTTCTATATAGTTTCAGCAATGGCCTGAATAGAAGCAACCAGGCACCAT  
CTCAGCAATGTTTTCTCTTGTTTGTAAATTATTTGCTCCTTTGAAAATTAAATCACTATTAATT  
ACATTAATAAATCAAATTGGATAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 452**

MTSKFILVSFILAAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVKQ  
VTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRINKSFSLDHMNIYDSKFWEEATPIWIT  
NQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLYWED  
PDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEERLIEL  
DQYLDKDHYTLLIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYNRIQP  
IIA VADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTDLYP LLC  
HLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPPYFIGVSLGS  
IIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

**Important features:****Signal Peptide:**

amino acids 1-22

**Transmembrane Domain:**

amino acids 429-452

**N-glycosylation sites:**

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372,  
382-385, 389-392

**Somatomedin B Domain:**

amino acids 69-85

**Sulfatase protein Region:**

amino acids 212-241



## FIGURE 453

GGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAACAGCTTGCTGAAGGAGCTGAGA**TG**  
**A**GGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTGAGGCTGAGGGGTGCCAGCGTGGGTG  
GGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCTTGGCCATTAAATGAAACCTGAAGG  
CCAAA  
AAA

**FIGURE 454**

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPKESSSTKETERKETKAEEEELD  
 AEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTNTY  
 TSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMVRLI  
 NKFNSSSSSLEEKIAALFDLEYVHQM DNAQDLLSFGGLQVVINGLNSTEPLVKEYAAAFVLGA  
 AFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKLGGLQV  
 LRTL VQEKGTEVLAVRVVTL LYDLVTEKMF AEEEEAE L TQEMSPEKLQOYRQVHLLPGLWEQGW  
 CEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLASLELQDGE  
 DEGYFQELLGSVNSLLKELR

**Important features:****Signal peptide:**

amino acids 1-29

**Hypothetical YJL126w/YLR351c/yhcX family protein.**

amino acids 364-373

**N-glycosylation site.**

amino acids 193-197, 236-240

**N-myristoylation site.**

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

**Homologous region SLS1 protein.**

amino acids 68-340

**FIGURE 455**

CCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCCGGTGCCCAGAGCCCAGGAGGAGGCAGT  
GGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCCTA  
CCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGGTGT  
CTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCTAGAC  
TCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGTCCTTG  
TGGTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCC**ATG**GCTACAGCAAGACCCCCCTGGA  
TGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTTCTCGCCA  
ACAATGATGTTTTCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCAGGACCTGG  
GAGCTGGGGCCGGGGAAGACGCCCGGTGCGATGACAGCAGCAGCCGCATCATCAATGGATCCG  
ACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAACCAGCTCTACT  
GCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAGGAAGAAAGTTT  
TCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGCAGCAGATGTTCC  
AGGGGGTCAAATCCATCCCCACCCTGGCTACTCCCACCCTGGCCACTCTAACGACCTCATGC  
TCATCAAACCTGAACAGAAGAATTCGTCCCCTAAAGATGTCAGACCCATCAACGTCTCCTCTC  
ATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAACCAAGAGCCCCCAAG  
TGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGTCAGAAAAGGTGCGAGG  
ATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAGAGACT  
CCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTT  
GGGGAGATTACCCTTGTGCCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTCA  
CCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCATCCCAGGACTCAGCACACCGG  
CATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTTACAGACCTCATTCCTTCCCAGAGA  
TGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCTCCTGGACTCAGGGTCTGCTTCC  
CCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGGGAACAATTTCCAAAACCTGTCCAG  
GGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAGCTCAGGGCCCATCCCTT  
CTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAACTGAGAAGTGGAACCAAAAAA

456/550

## **FIGURE 456**

MATARPPWMWVLCALITALLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDDS  
SSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLSPV  
YESGQQMFQGVKSIHPHGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCP SAGTKCLVSG  
WGTTKSPQVHF PKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGR DSCQGD SGGPVVCN  
GSLQGLVSWGDYPCARPNRPGVYT NLCKFTKWIQETIQANS

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted April 28, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**FIGURE 457**

GCAGTCAGAGACTTCCCCTGCCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCTTGCTTCCTGAACT  
 AGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCTCACCGCTGTAGGAATCCAGATGCAGG  
 CCAAGTACAGCAGCACGAGGGACATGCTGGATGATGATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTG  
 CCACAACTCGGCATCCAGAGCCCCGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAAGTGGCCCTGA  
 CCCTGCTGACTTTGTGCTTGGTGTCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACCAGC  
 TCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCCCCAAGAGTTGCAATCTC  
 TTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGCTGAAAACTCTGTGCTGAGCTGTATAACA  
 AAGCTGGAGCACACAGGTGCAGCCCTTGTACAGAACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATA  
 AAGACAGCAAAAGTTGGGAGGACTGTAAATATTTCTGCCTTAGTGAAAACTCTACCATGCTGAAGATAAACAAAC  
 AAGAAGACCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTTTGGCGC  
 CTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACTTCTGAACTGTTCCATATTATAATAGATG  
 TCACCAGCCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGGGATGATCTTCTCAAAGGACTGCAAAGAATTGA  
 AGCGTTGTGTCTGTGAGAGAAGGGCAGGAATGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCG  
 AAGGTGACTTGATTTCGCCCTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAG  
 ACATTGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCCTGTTCTCTGTT  
 CAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACAAGAAGTCTTATTTACATGCCAC  
 CAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTGGCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAA  
 TGTCTAATATCACCTCCCTGTTTTTCATGTCTTCCTTACACTTGGTGAATAAGAACTTTTTGAAGTAGAGGAAA  
 TACATTGAGGTAACATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTAC  
 CAGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCATCAGTAAAGACCC  
 CATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAGAATCTCAAATCTCAATGCCTTATAA  
 GCATTCTTCTGTGTCCATTAAGACTCTGATAATTGTCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATAT  
 ATCCCCATCTCCGTTTTCATATCAGAACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAGT  
 GAGCCTCTTCATCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAGGT  
 ACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

458/550

**FIGURE 458**

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLLI  
GLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRELY  
NKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAASQSY  
SEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKDCKEL  
KRCVCERRAGMVKPESLHVPPETLGED

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 459**

GTTGATGGCAAACCTTCCTCAAAGGAGGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCAC  
 TGGCGGCCCCGCAACACTCCGTCTCACCTCTGGGCCCCACTGCATCTAGAGGAGGGCCGTCTGT  
 GAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGTGGT  
 CAGCTGGGTGAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAGGGAA  
 GTGGGAGCCTCGAGCCCTCGGGTGGAAGCTGACCCCAAGCCACCCTTCACCTGGACAGG**ATGA**  
 GAGTGTGAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGATGTTTA  
 TTCGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCAGCCTCGC  
 CCACCAAGGAGATCCAGGTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCCCAGCCAACT  
 ACTTTGCGTTTAAAATCTGCAGTGGGGCCGCCAACGTCTGGGCCCTACTATGTGCTTTGAAG  
 ACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATCGCCCTGGTGA  
 ATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGGAGATGTTATGC  
 ACCTAGTGAAATTCCTTAAAGAAATTCGGGGGGTGCCTGGTGCTGGTGGCCTCCTACGACG  
 ATCCAGGGACCAAAATGAACGATGAAAGCAGGAAACTCTTCTCTGACTTGGGGAGTTCCTACG  
 CAAAACAACTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGACCTCAGGGGTAAAA  
 GCCCCTTTGAGCAGTTCTTAAAGAACAGCCAGACACAAACAAATACGAGGGATGGCCAGAGC  
 TGCTGGAGATGGAGGGCTGCATGCCCCGAAGCCATTT**TAG**GGTGGCTGTGGCTCTTCCTCAG  
 CCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCGGCAGGGGCTGAGGAGGA  
 GGAGCAGGGGGTGTGCTGCGTGGAAGGTGCTGCAGGTCCTTGACGCTGTGTGCGCCTCTCCTC  
 CTCGGAAACAGAACCCTCCCACAGCACATCCTACCCGGAAGACCAGCCTCAGAGGGTCCTTCT  
 GGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTCGTGAGGGACTGCTGACGGCTGGTCCTG  
 AGGAAGGACAACTGCCCAGACTTGAGCCCAATTAAATTTTATTTTGGCTGGTTTTGAAAAAA  
 AAAAAAAAAAAAAA

## **FIGURE 460**

MRVSGVLRLLALIFAIVTTWMFIRSYMSFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCPA  
NYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSGDV  
MHLVKFLKEIPGGALVLVASYYDDPGTKMNDESRKLFSDLGSSYAKQLGFRDSWVFIGAKDLRG  
KSPFEQFLKNSPDTNKEYEGWPELLEMEGCMPPKPF

### **Important features:**

#### **Signal peptide:**

amino acids 1-15

#### **ATP/GTP-binding site motif A (P-loop).**

amino acids 184-191

#### **N-glycosylation site.**

amino acids 107-110



**FIGURE 461**

AACCTCAGCACTTGCCGGAGTGGCTCATTGTTTAAGACAAAGGGTGTGCACCTTCCTGGCCAGGA  
 AACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCTCAGAACCTTCCTC  
 TAGAACCCGACCCACCACC**ATG**AGGTCTCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGGCGT  
 CCAGTGGTCTTTGCTTCTGGCTGTCCTGGTCTTCTTTCTCTTCGCCTTGCCCTCTTTTATTAA  
 GGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCTCTACA  
 GTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTATGCAGA  
 GCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCCAAGGCCACACCACCGGAG  
 CAGAGGAAAGGAGGCCAACCCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCACACAGCACA  
 GAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAAACCATGGTGAACACACTGTCACCCAGAGG  
 GCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCCAGGACACAAA  
 GACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACGGTGTGAGAGAA  
 GCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTTCCAAAAGTCAGCACAGAATGCT  
 GGCTCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCACAGCAGTCATCCC  
 ACCTAAGGAGAAGAAACCTCAGGCCACCCCACCCCTGCCCTTTCCAGAGCCCCACGACGCA  
 GAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTGGGATTTTGAGGAAAA  
 ATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTGTGAAGATCAAAGCCTC  
 CAAGTCGCTGTGGCTCCAGAAACTCTTTCTGCCAACCTCACTCTCTTCTGGACTCCAGACA  
 CTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACCCCTTTGGCTTCATGGAGCT  
 CAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAGTGCCCCAGCAGCAGCTGCT  
 CCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGTGCCGTGGTGGGCAACGGGGG  
 CATCCTGAACAACCTCCACATGGGCCAGGAGATAGACAGTCACGACTACGTGTTCCGATTGAG  
 CGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTCGGACATCCTTCTACGGCTTTAC  
 CGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAATCGGGGTTTCAAGAACGTGCCTCT  
 TGGGAAGGACGTCCGCTACTTGCACTTTCCTGGAAGGCACCCGGGACTATGAGTGGCTGGAAGC  
 ACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTTTTCTGGTTCAGGCACAGACCCCAGGA  
 AGCTTTTTCGGGAAGCCCTGCACATGGACAGGTACCTGTTGCTGCACCCAGACTTTTCTCCGATA  
 CATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGATGGTGGCCACTGGAGGATATACCGCCC  
 CACCCTGGGGCCCTCCTGCTGCTCACTGCCCTTCAGCTCTGTGACCAGGTGAGTGCTTATGG  
 CTTTCACTACTGAGGGCCATGAGCGCTTTTCTGATCACTACTATGATACATCATGGAAGCGGCT  
 GATCTTTTACATAAACCATGACTTCAAGCTGGAGAGAGAAGTCTGGAAGCGGCTACACGATGA  
 AGGGATAATCCGGCTGTACCAGCGTCTGGTCCCAGGAACTGCCAAAGCCAAGAACT**TGA**CCGGG  
 GCCAGGGCTGCCATGGTCTCCTTGCTTCAAGGCACAGGATACAGTGGGAATCTTGAGAC  
 TCTTTGGCCATTTCCCATGGCTCAGACTAAGCTCCAAGCCCTTCAGGAGTTCCAAGGGAAACAC  
 TTGAACCATGGACAAGACTCTCTCAAGATGGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCA  
 GTACATTGCTGTAGGTCTTGAGGCCAGGGATTTTTTAATTAAATGGGGTGATGGGTGGCCAATA  
 CCACAATTCTGCTGAAAAACACTCTTCCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCC  
 TGGATTTACAGAAACATATAGATCTGGTTTGAATTCCAGATCGAGTTACAGTTGTGAAATCT  
 TGAAGGTATTACTTAACTTCACTACAGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTC  
 TAGAAGGGTCTATACTTGTCTTGTCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAAC  
 TGATAATAATACAAATGATTGTTGTCCATGGAAAGGCAAATAAATTTTCTACAGTGAAAAAAA  
 AAAAAAAA

**FIGURE 462**

MRSCLWRCRHLSSQGVQWSLLLVFLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKPK  
SQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAWKS  
PEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQGKAA  
TTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQORNQRLK  
AANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHFNQSEW  
DRLEHFAPPFGFMELNYSLVQKVVTFRFPVPQQQLLLASLPAGSLRCITCAVVGNNGGILNNSH  
MGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFATAFSLTQSLILGNRGFKNVPLGKDVR  
LHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFFREALHMDRYLLLHPDFLRYMKNRFL  
RSKTLDDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSWKRLIFYINH  
DFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

**Important features:****Cytoplasmic Domain:**

amino acids 1-10

**Type II Transmembrane Domain:**

amino acids 11-35

**Lumenal catalytic Domain:**

amino acids 36-600

**Ribonucleotide Reductase small subunit Signature:**

amino acids 481-496

**N-glycosylation Sites:**

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGCGGCCGCAAGGGTGAGGGCGGCCCCAGAAC  
CCCAGGTAGGTAGAGCAAGAAG**ATGGT**GTTTCTGCCCTCAAATGGTCCCTTGCAACCATGTCT  
ATTTCTACTTTTCTCACTGTTGGCTCTCTTAACCTGTGTCCACTCCTTCATGGTGTCTAGAGCAC  
TGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCCTTGGAAATAAAATACGACTTCTCTGA  
GTACGTATCCCCAGTTTCATTATGATCTCTTGATCCATGCAAACTTACCACGCTGACCTTCTGT  
GGGAACACGAAAGTAGAAATCACAGCCAGTCAGCCCAACGACCACTCATCTCTGCATAGTCA  
CCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGGCTATCGGAAGAACC  
CCTGCAAGTCTCTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAGCCCCCTCCT  
TGTCGGGCTCCCCGATACAGAGTTGTCATTCATATGCTGGCAATCTTTCGGAGACTTTCACCG  
ATTTTACAAAAGCACCTACAGAACCAAGGAAGGGAACTGAGGATACATGACATCAACAAAT  
TGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGATGAACCTGCCTTCAAAGCAAGTTT  
CTCAATCAAATTAGAAGAGAGCCAAGGCACCTAGCCATCTCCAATATGCCATTGGTGAAATC  
TGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGTATGTCACTGTGAAGATGAGCACCTA  
TCTGGTGGCCTTCATCTTTCAGATTTTGGATCTGTGCAGCAAGATAACCAAGAGTGGAGTCAA  
GGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGCATCGGATGCTGCGGT  
GACTCTTCTAGAATTTTATGAGGATTATTTTCAGCATACCGTATCCCCCTACCCAAACAAGATCT  
TGCTGCTATTCCCAGCTTTCAGTCTGGTGTCTATGAAAACCTGGGGACTGACAACATATAGAGA  
ATCTGCTCTGTTGTTTGTATGCAGAAAAGTCTTCTGCATCAAGTAAGCTTGGCATCACAGTGAC  
TGTGGCCCATGAACGTGGCCACCAAGTGGTTTGGGAACCTGGTCACTATGGAATGGTGGAAATGA  
TCTTTGGCTAAATGAAGGATTGCCAAATTTATGGAGTTTGTGTCTGTCACTGTGACCCATCC  
TGAAGTGAAGGTTGGAGATTATTTCTTTGGCAAATGTTTTGACGCAATGGAGGTAGATGCTTT  
AAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTGCTCAGATCCGGGAGATGTTTGA  
TGATGTTTCTTATGATAAGGGAGCTTGTATTTCTGAATATGCTAAGGGAGTATCTTAGCGCTGA  
CGCATTTAAAAGTGGTATTGTACAGTATCTCCAGAAGCATAGCTATAAAAATACAAAAACGA  
GGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAGATGGTGTAAAGGGATGGATGGCTT  
TTGCTCTAGAAGTCAACATTTCATCTTCATCCTCACATTGGCATCAGGAAGGGGTGGATGTGAA  
AACCATGATGAACATTTGGACACTGCAGAGGGGTTTCCCCCTAATAACCATCACAGTGAAGGG  
GAGGAATGTACACATGAAGCAAGAGCTACATGAAGGGCTGTGACGGCGCCCCGGACACTGG  
GTACCTGTGGCATGTTCCATTGACATTCATCACCAGCAAATCCAACATGGTCCATCGATTTTT  
GCTAAAAACAAAACAGATGTGCTCATCCTCCCAGAAGAGGTGGAATGGATCAAATTTAATGT  
GGGCATGAATGGCTATTACATTGTGCATTACGAGGATGATGGATGGGACTCTTTGACTGGCCT  
TTTAAAGGAACACACACAGCAGTCAAGCATTAATGATCGGGCAAGTCTCATTAAACAATGCATTT  
TCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCCTTGGATTTATCCCTGTACTTGAA  
ACATGAAACTGAAATTATGCCCGTGTTTCAAGGTTTGAATGAGCTGATTCTCATGTATAAGTT  
AATGGAGAAAAGAGATATGAATGAAGTGGAAACTCAATTCAAGGCCTTCTCATCAGGCTGTCT  
AAGGGACCTCATTGATAAGCAGACATGGACAGACGAGGGCTCAGTCTCAGAGCAAATGCTGCG  
GAGTGAACCTACTACTCCTCGCTGTGTGCACAACTATCAGCCGTGCGTACAGAGGGCGAAGG  
CTATTTTCAGAAAGTGAAGGAATCCAATTGAAACTTGAGCCTGCCTGTGCGAGCTGACCTTGGC  
AGTGTTTGCTGTGGGGGCCCAGAGCACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTT  
TTCTTTGTCCAGTACTGAGAAAAGCCAAATTGAATTTGCCCTCTGCAGAACCCAAAATAAGGA  
AAAGCTTCAATGGCTACTAGATGAAAGCTTTAAGGGAGATAAAATAAAAACCTCAGGAGTTTCC  
ACAAATTTTACACTCATTGGCAGGAACCCAGTAGGATACCCACTGGCCTGGCAATTTTCTGAG  
GAAAAACTGGAACAACTTGTACAAAAGTTTGAACCTGGCTCATCTTCCATAGCCACATGGT  
AATGGGTACAACAAATCAATTCTCCACAAGAACACGGCTTGAAGAGGTAAAAGGATTCTTCAG  
CTCTTTGAAAGAAAATGGTTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATTGAAGA  
AAACATCGTGTGGATGGATAAGAATTTTGATAAAATCAGAGTGTTGGCTGCAAAGTGAAAAGCT  
TGAACGTATG**TAA**AAATTCCTCCCTTGCCCGTTTCTGTTATCTCTAATCAACCAACTTTTGT  
TGAGTGTATTTTCAAACCTAGAGATGGCTGTTTTGGCTCCAACCTGGAGATACTTTTTTCCCTTC  
AACTCATTTTTTTGACTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTTCATGAATGGGCTTTTT  
CATGAATGGGCTATCGCTACCATGTGTTTTGTTCATCACAGGTGTTGCCCTGCAACGTAAACC  
CAAGTGTGGTTCCCTGCCACAGAAGAATAAAGTACCTTATTTCTCAAAAAAAAAAAAAA  
AAAAAAAAAAAAA

**FIGURE 464**

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVHY  
 DLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLOVLEHP  
 PQEQIALLAPEPLLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTAARM  
 AFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVAFIIS  
 DFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLFEFYEDYFSIPYPLPKQDLAAIPDFQ  
 SGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDLWLNEGF  
 AKFMEFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFDDVSYDKG  
 ACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDGFCRSQS  
 SSSSHWHQEGVDVKTMMNWTWLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPDTGYLWHVPL  
 TFITSKSNMVHRFLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSLTGLLKGTHTA  
 VSSNDRASLINNAFQLVVSIGKLSIEKALDLSLYLKHETEIMPVVFQGLNELIPMYKLMEKRDMN  
 EVETQFKAFLIRLLRDLIDKQTTWTDEGSVSEQMLRSELLLLACVHNYQPCVQRAEGYFRKWKE  
 SNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSSTEKSQIEFALCRTQNKELQWLLD  
 ESFKGDKIKTQEFPPQILTILGRNPVGYPLAWQFLRKNWNKLVQKFELGSSSIAHMMVGTNQF  
 STRTRLEEVKGFFSSLKENGSQLRCVQQTITETIEENIGWMDKNFDKIRVWLQSEKLERM

**Important features:**

**Signal peptide:**

amino acids 1-34

**N-glycosylation sites:**

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

**Neutral zinc metallopeptidases, zinc-binding region signature:**

amino acids 350-360

**FIGURE 465**

CAGCCACAGACGGGTCATGAGCGGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCACT  
 GCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCCGA  
 CCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGACAC  
 GTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGGAGGC  
 CAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATCTCCTA  
 CACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCTTTGGGC  
 CCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTCTGCTTGTCTATGGAAGGCTG  
 TCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGATGGCCTCCT  
 CAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCCCCAGCCAGG  
 TTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGTGGGTATGACTGAGAACTGCAATAG  
 GAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAACTTGGCTCAAGA  
 ACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGTGTGTGAGGAGAC  
 GCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAGGCTGCAGCACTGT  
 TGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGGGTGCTTGTGGCCTC  
 CTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAGCAGCGTTCTGCTGAA  
 CTCCCTCCCTCCTCAAGCTGCCCCGTGCCAGGAGACCGGCAGTGTCTACCTGTGTGCAGCC  
 CTTTGGAAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGGGGCGCCACTCATTGTTA  
 TGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAATGAGCATTCAGGGCTGCGT  
 GGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCGGGATCTTCTCTGCGCGTGA  
 GAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGTGGGGCTGAGGGCCTGGAGTC  
 TCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTGGTGGGGAGTGGTTTGGCCCTC  
 CTGCTTAACTCTATTACCCCCACGATTCCTTACCGCTGCTGACCACCCACACTCAACCTCCCTC  
 TGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTCCATTCTGTCCATGAATCATCTT  
 CCCCACACACAATCATTCATATCTACTCACCTAACAGCAACACTGGGGAGAGCCTGGAGCATC  
 CGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTGGCTGCATGTATCTGATAATACAGAC  
 CCTGTCCTTTCA

**FIGURE 466**

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLIE  
SGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQPPA  
DPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPPQPGCNLLN  
GTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETLLLD  
VGLTSTLVGKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLNSLPPQ  
AAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGCVAQPSS  
FLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVCPSC

[illegible]

**FIGURE 468**

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQEF  
SKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLLQE  
AEEKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSGWMDAVYNSPKVYLLIGSRNNTVWEF  
ANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVEDRMLL  
PGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGLGVEHSWDTPCRSQ  
DAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIHYNPRDK  
QLYAWNEGNQIIYKLQTKRKLPLK



**FIGURE 469**

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGGC  
AGTGTTTTGCCTTCACCCCAAGTGACCATGAGAGGTGCCACGCGAGTCTCAATCATGCTCCTC  
CTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCTGTGAGCGGGATGTCCAGTGTGGG  
GCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCTGGGG  
CGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCAAGCAC  
CACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTACCGCTGC  
TCCATGGACTTGAAGAACATCAATTTTTAGGCCTTGCTGCTCAGGATACCCACCATCCT  
TTTCCTGAGCACAGCCTGGATTTTTTATTTCTGCCATGAAACCCAGCTCCCATGACTCTCCAG  
TCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAGGCAGACATA  
CCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGGCTGTGGTGTG  
AAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTAAATGGCAGAAA  
GGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCTGGGCCCTGCCCCTCTCCCCA  
CATGTATCCCTCGGTCTGAATTAGACATTCCCTGGGCACAGGCTCTTGGGTGCATTGCTCAGAG  
TCCAGGTCTTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGGACCAATTTGTGGGT  
AGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGACTCAAGATTGGCTCTT  
CCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCAGGGAGGCCAATCAGCC  
CCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGACCTGTGACCTTCTGCCA  
GAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGTTAACCCTGAAGCCCCCA  
ATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAATCTAATCTGATATTGACAT  
ATTAGAAGGCAATTAGGGTGTTCCTTAAACAACCTCCTTTCCAAGGATCAGCCCTGAGAGCAG  
GTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGGTGGGAGCAAGGGACAGGGAGC  
AGGGCAGGGGCTGAAAGGGGCACTGATTTCAGACCAGGGAGGCAACTACACACCAACATGCTGG  
CTTTAGAATAAAAGCACCAACTGAAAAA

## **FIGURE 470**

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEETCHPG  
SHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

### **Important features:**

#### **Signal peptide:**

amino acids 1-19

#### **Tyrosine kinase phosphorylation site:**

amino acids 88-95

#### **N-myristoylation sites:**

amino acids 33-39, 35-41, 46-52

**FIGURE 471**

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAAATGTCTTTCC  
TCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGCTG  
CTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGGCCC  
TGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAAGCAA  
AGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTTCCTCT  
GTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCGTCCCCC  
TCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTATTTCAAAG  
GAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTCCACAAAGGCGGAAGATGATGTTTA  
TGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCTTCCGAGCCTGGAACGGAGGCTTCTCTG  
GAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTTCGTGGTGGGATCAGGAAAGCAGG  
GCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACTTTCTGTTCTGG  
AAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAATGATTGTGTGAAACTG  
CCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTTATGGGATGTATTGTTTCCACTCGTG  
TCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTATTAATGTATTTTAA  
TATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGACAAAAATCTGAAAAA  
CTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAATTGACTGCCAGGCTGG  
GTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGAGCAAGTCACTTGAG  
GTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAA  
ATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCGGGAGGCTGAGGCAGGAGAA  
TCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCACACCACTGTATTCCAGCCTG  
GGTGACTGAGACTCTAACTAA

472/550

**FIGURE 472**

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKALEYLEDIDLKTLEKEPRT  
FKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDFQP  
YFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEFGFILGGVFVVG  
GKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

472/550

**FIGURE 473**

AATATATCATCTATTTATCATTAAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTTT  
TGGGATTTTAATTTTCAAACACAGCAGA**ATG**ACATTTTTCTGTCACTATTATTATTGTTGGT  
ATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCTATC  
AAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTCCTAC  
CAAAGCTGTCAAACACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGAGGGTT  
AATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACT**TAG**ATTCAATGATTGTA  
AATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATATATACAA  
TAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACTTTATTAAT  
TTTTAAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGATCATATAAT  
TTGATACAAATAAAAGAAAAGTGTTCTCTCCCCTTACAGAATTGACATTTTAAATGCGATACA  
GTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAAAGAAGGGAAAA  
TGTTGCCAAGGAAAAAAAAA

474/550

**FIGURE 474**

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGKG  
IVKGRNLDSRGLILGAEAWGRGVKKNT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 476**

MAPLALHLLVLVPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVVVTWGLNRTLKPQRVI  
VVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRILHK  
LCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGYPALRPQEKGHSPEDIYQMALNQ  
ALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSDGGFFYLSFAEALRAHSC  
LSDRLQYSRIVGGWDLPLRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNLKVLKA  
DVVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGGHSNTDR  
PSRMI FYPPPREGALLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQLWDGTGVV  
KRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKSALRAAIKI  
NSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQNTTHTRTSH

**Important features:**

**Signal peptide:**

amino acids 1-21





**FIGURE 478**

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFKLLFTKMLFIFNF  
 LFSPLPTPALICILTFGAAIFLWLITRPQPVLPDLLNNSVGIIEGGARKGVSQKNNDLTSCC  
 FSDAKTMYEVFQRLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSSPDQ  
 FVGIFAQNRPEWIISELACYTYSMAVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKALVLIG  
 NVEKGFTPSLKVIIILMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPPSPEDLSVIC  
 FTSGTTGDPKGAMITHQNIVSNAAAFKLCVEHAYEPTPDDVAISYLPLAHMFERIVQAVVYSC  
 GARVGFFQGDIRLLADDMKTLKPTLFPAPVPRLLNRIYDKVQNEAKTPLKKFLLKLAVSSKFKE  
 LQKGIIIRHDSFWDKLIIFAKIQDSLGGRRVRVIVTGAAPMSTSVMTFFRAAMGCQVYEAYGQTEC  
 TGGCTFTLPGDWTSGHVGVPACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFKGYLKDPEKT  
 QEALDSDGWLHTGDIGRWLPNGTLKIIDRKKNIFKLAQGEYIAPEKIENIYNRSQPVLQIFVH  
 GESLRSSLVGVVVPDTPDLPSFAAKLGVKGSFEELCQNQVVRREILEDLQKIGKESGLKTFEQ  
 VKAIFLHPEPFPSIENGLLTPTLKAKRGELSKYFRTQIDSLYEHIQD

**Important features:**

**Type II transmembrane domain:**

amino acids 61-80

**Putative AMP-binding domain signature.**

amino acids 314-325

**N-glycosylation site.**

amino acids 102-105, 588-591 and 619-622

**FIGURE 479**

GGAGGCGGAGGCCGCGGCGAGCCGGGCCGAGCAGTGAGGGGCCCTAGCGGGGCCCCGAGCGGGGC  
CCGGGGCCCCCTAAGCCATTTCCTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAATCC  
GGT**ATG**GACGACTGGAAGCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGGAGC  
TGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGCCCTGCGGAGATTCTGTCAGACA  
GGGGCCGTGCTTTTCCTGCTGGTGAAGTGTCAATATCAAGTTGATCCTGGACACTCGG  
CGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGAGCAAGACTATGATGAGGCCCTAGGC  
CGCCTGGAGCCCCCACGGGCGCAGAGGCAGTGGTCCCCGGCGGGTCTGAGCGTAGAGGTGTAT  
TCAAGTCGCAGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGCTGGAGGATGAGGCCCGG  
GAGCAGGGCCCGGGGCATCCATGTCATTGTCCTCAACCAGGCCACGGGCCACGTGATGGCAAAA  
CGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGCTATTTCCTCAACATGGTA  
GCGCCCGGGCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACCTCAAGGACACA  
GCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGCTGGAGGGACACA  
TGGGCCTTCGTGGGACGAAAAGGAGGTCTGTCTTCGGGGAGAAACATTCTAAGTCACCTGCC  
CTCTCTTCCTGGGGGGACCCAGTCCTGCTGAAGACAGATGTGCCATTGAGCTCAGCAGAAGAG  
GCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGCTTCTGCAGCAAAGTT  
GAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCACACCCATCGAGTTCAGCCCTGACCCA  
CTCCCAGACAACAAGGTCCTCAATGTGCCTGTGGCTGTCAATTGCAGGGAACCGACCCAATTAC  
CTGTACAGGATGCTGCGCTCTCTGCTTTTCAGCCCAGGGGGTGTCTCCTCAGATGATAACAGTT  
TTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGGCACTGTTTGGTCTGAGGGGCATC  
CAGCATACTCCCATCAGCATCAAGAATGCCCCGCGTGTCTCAGCACTACAAGGCCAGCCTCACT  
GCCACTTTC AACCTGTTTCCGGAGGCCAAGTTTGCTGTGGTTCTGGAAGAGGACCTGGACATT  
GCTGTGGATTTTTTTCAGTTTTCCTGAGCCAATCCATCCACCTACTGGAGGAGGATGACAGCCTG  
TACTGCATCTCTGCCTGGAATGACCAGGGGTATGAACACACGGCTGAGGACCCAGCACTACTG  
TACCGTGTGGAGACCATGCCTGGGCTGGGCTGGGTGCTCAGGAGGTCCTTGTACAAGGAGGAG  
CTTGAGCCCCAAGTGGCCTACACCGGAAAAGCTCTGGGATTGGGACATGTGGATGCGGATGCCT  
GAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACGTTTCCCGATCCTACCACTTTGGCATC  
GTCGGCCTCAACATGAATGGCTACTTTACAGAGGCCTACTTCAAGAAGCACAAGTTCAACACG  
GTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCTGAAGAAAGAAGCTTATGAAGTGAAGTT  
CACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACCACAGCAAGAACCCTTGTGAAGACTCTTTC  
CTGCCAGACACAGAGGGCCACACCTACGTGGCCTTTTATTGCAATGGAGAAAGATGATGACTTC  
ACCACCTGGACCCAGCTTGCCAAGTGCCTCCATATCTGGGACCTGGATGTGCGTGGCAACCAT  
CGGGGCCTGTGGAGATTGTTTTCGGAAGAAGAACCCTTCCTGGTGGTGGGGTCCCGGCTTCC  
CCCTACTCAGTGAAGAAGCCACCCCTCAGTCACCCCAATTTTCTGGAGCCACCCCCAAAGGAG  
GAGGGAGCCCCAGGAGCCCCAGAACAGACAT**TGA**GACCTCCTCCAGGACCCTGCGGGGCTGGGT  
ACTGTGTACCCCCAGGCTGGCTAGCCCTTCCCTCCATCCTGTAGGATTTTGTAGATGCTGGTA  
GGGGCTGGGGCTACCTTGTTTTTAAACATGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCC  
CTGCTCCAACACCCCGTTCCCTGAGTTAAAAGTCTATTTATTTACTTCTTGTGGAGAAGGGC  
AGGAGAGTACCTGGGAATCATTACGATCCCTAGCAGCTCATCCTGCCCTTTGAATACCCTCAC  
TTTCCAGGCCTGGCTCAGAATCTAACCTATTTATTGACTGTCCTGAGGGCCTTGAAAACAGGC  
CGAACCTGGAGGGCCTGGATTTCTTTTTGGGCTGGAATGCTGCCCTGAGGGTGGGGCTGGCTC  
TTACTCAGGAACTGCTGTGCCCAACCCATGGACAGGCCAGCTGGGGCCCACATGCTGACAC  
AGACTCACTCAGAGACCCTTAGACACTGGACCAGGCCTCCTCTCAGCCTTCTCTTGTCCAGA  
TTTCCAAAGCTGGATAAGTTGGTCATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAA  
AAAAAAAAAAAAAAAAAA

**FIGURE 480**

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQALRRFCQTGAVLFLLVTVIVNIKLILDTRR  
 AISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEARE  
 QGRGIHVIVLNQATGHVMAKRVFDTYSPHEDEAMVLFILNMVAPGRVLICTVKDEGSFHLKDTA  
 KALLRSLGSQAGPALGWRDWTAFVGRKGPGVFGEKHSKSPALSSWGDPVLLKTDVPLSSAEEA  
 ECHWADTELNRRRRRRFC SKVEGYGSVCCKDPTPIEFSPDPLPDNKVLNVPVAVIAGNRPNYL  
 YRMLRSLLSAQGVSPQMITVVFIDGYEPMDDVVALFGLRGIQHTPISIKNARVSQHYKASLTA  
 TFNLFPEAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSLYCISAWNDQGYEHTAEDPALLY  
 RVETMPGLGWVLRSLYKEELEPKWPTPEKLWDWDMWMMRMPEQRRGRECIIPDVSRSYHFGIV  
 GLNMNGYFHEAYFKKHKFNTVPGVQLRNVDLKEAYEVEVHRLLEAEVLDHKNPCEDSFL  
 PDTEGHTYVAFIRMEKDDDDFTTWTQLAKCLHIWDLVVRGNHRGLWRLFRKKNHFLVVGVPASP  
 YSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

**Important features:**

**Transmembrane domain:**

amino acids 38-55

**Homologous region to Mouse GNT1**

amino acids 229-660

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 481**

GAAAGAATGTTGTGGCTGCTCTTTTTTCTGGTGA CTGCCATTCATGCTGAACTCTGTCAACCA  
GGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCATAT  
GCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAAGTT  
CCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGTATCA  
TTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGCAATCA  
GCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAA ACTCTGGAA  
TTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTGGATTATT  
ATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTATCAGGGATC  
TGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGATAAGTGTGAA  
AACATGATCACAATTGAAAATGGCATCCCCCTCTGATCCCCCTGGACATGAAGGGGGGCATATTA  
ATGATGCCTTCATGACAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGTTGTTCTGCTTC  
CTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATACCAAGAGCAGAT  
CATATATTTTGTTCACCATTCTTCTTTTGTAATAAATTTTGAATGTGCTTGAAAGTGAAGAG  
CAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGACTCAAATATTTCTAA  
AATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTGTAGTTATTGATTTAA  
GCATTTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTCAAAGACCTAAGGAAAA  
ATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCATTGAAAATGGATCCTTTT  
TGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAGTGAGAAGTAATTATTGTA  
AATGGATGGATAAAAATGGAATTACTCATATACAGGGTGAATTTTATCCTGTTATCACACCA  
ACAGTTGATTATATATTTTCTGAATATCAGCCCCTAATAGGACAATTCTATTTGTTGACCATT  
TCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAGTAATAATCATCTCTTTTTTAA  
AAAAAAAAAAAAAAAAAAAAA

**FIGURE 482**

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRLTALGDKAYAWDTNEEYLFKAMVAFSMRKVPN  
REATEISHVLLCNVTQRVSWFVVTDP SKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLEFL  
KIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKCENM  
ITIENGIPSDPLDMKGGILMMPS

**FIGURE 483**

CGTCTCTGCGTTTCGCC**ATG**CGTCCCGGGGCGCCAGGGCCACTCTGGCCTCTGCCCTGGGGGGC  
CCTGGCTTGGGCCGTGGGCTTCGTGAGCTCCATGGGCTCGGGGAACCCCGCGCCCGGTGGTGT  
TTGCTGGCTCCAGCAGGGCCAGGAGGCCACCTGCAGCCTGGTGCTCCAGACTGATGTCACCCG  
GGCCGAGTGCTGTGCCTCCGGCAACATTGACACCGCCTGGTCCAACCTCACCCACCCGGGGAA  
CAAGATCAACCTCCTCGGCTTCTTGGGCCTTGTCCTACTGCCTTCCCTGCAAAGATTTCGTGCGA  
CGGCGTGGAGTGCGGCCCGGGCAAGGCGTGCCGCATGCTGGGGGGCGCCCGCGCTGCGAGTG  
CGCGCCCGACTGCTCGGGGCTCCCGGCGCGGCTGCAGGTCTGCGGCTCAGACGGCGCCACCTA  
CCGCGACGAGTGCGAGCTGCGCGCCGCGCGCTGCCGCGGCCACCCGGACCTGAGCGTCATGTA  
CCGGGGCCGCTGCCGCAAGTCCTGTGAGCACGTGGTGTGCCCCGCGGCCACAGTCGTGCGTCGT  
GGACCAGACGGGCAGCGCCCACTGCGTGGTGTGTGAGCGGCGCCCTGCCCTGTGCCCTCCAG  
CCCCGGCCAGGAGCTTTGCGGCAACAACAACGTCACCTACATCTCCTCGTGCCACATGCGCCA  
GGCCACCTGCTTCTTGGGGCGCTCCATCGGCGTGCGCCACGCGGGCAGCTGCGCAGGCACCCC  
TGAGGAGCCGCCAGGTGGTGAGTCTGCAGAAGAGGAAGAGAACTTCGT**GTG**AGCCTGCAGGAC  
AGGCCTGGGCCTGGTGGCCGAGGCCCCCATCATCCCCTGTTATTTATTGCCACAGCAGAGTC  
TAATTTATATGCCACGGACACTCCTTAGAGCCCGGATTTCGGACCACTTGGGGATCCCAGAACC  
TCCCTGACGATATCCTGGAAGGACTGAGGAAGGGAGGCCTGGGGGCGGCTGGTGGGTGGGAT  
AGACCTGCGTTCGGGACACTGAGCGCCTGATTTAGGGCCCTTCTCTAGGATGCCCCAGCCCCCT  
ACCCTAAGACCTATTGCCGGGGAGGATTCCACACTTCCGCTCCTTTGGGGATAAACCTATTAA  
TTATTGCTACTATCAAGAGGGCTGGGCATTCTCTGCTGGTAATTCCTGAAGAGGCATGACTGC  
TTTTCTCAGCCCCAAGCCTCTAGTCTGGGTGTGTACGGAGGGTCTAGCCTGGGTGTGTACGGA  
GGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGT  
GAGTACGGAGGGTCTAGCCTGGGTGTGTATGGAGGATCTAGCCTGGGTGAGTATGGAGGGTCT  
AGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTGGGTGAGTAT  
GGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTG  
GGTGTGTACGGAGGGTCTAGTCTGAGTGCCTGTGGGGACCTCAGAACACTGTGACCTTAGCCC  
AGCAAGCCAGGCCCTTCATGAAGGCCAAGAAGGCTGCCACCATTCCCTGCCAGCCCAAGAACT  
CCAGCTTCCCCACTGCCTCTGTGTGCCCTTTGCGTCCTGTGAAGGCCATTGAGAAATGCCCA  
GTGTGCCCCCTGGGAAAGGGCACGGCCTGTGCTCCTGACACGGGCTGTGCTTGGCCACAGAAC  
CACCCAGCGTCTCCCCTGCTGCTGTCCACGTCAGTTCATGAGGCAACGTCGCGTGGTCTCAGA  
CGTGGAGCAGCCAGCGGACGCTCAGAGCAGGGCACTGTGTCCGGCGGAGCCAAGTCCACTCTG  
GGGGAGCTCTGGCGGGGACCACGGGCCACTGCTCACCCACTGGCCCCGAGGGGGGTGTAGACG  
CCAAGACTCACGCATGTGTGACATCCGGAGTCCTGGAGCCGGGTGTCCCAGTGGCACCCTAG  
GTGCCTGCTGCCTCCACAGTGGGGTTCACACCCAGGGCTCCTTGGTCCCCCACAACCTGCCCC  
GGCCAGGCCTGCAGACCCAGACTCCAGCCAGACCTGCCTCACCCACCAATGCAGCCGGGGCTG  
GCGACACCAGCCAGGTGCTGGTCTTGGGCCAGTTCTCCACGACGGCTCACCCCTCCCCTCCAT  
CTGCGTTGATGCTCAGAATCGCCTACCTGTGCCTGCGTGTAACCACAGCCTCAGACCAGCTA  
TGGGGAGAGGACAACACGGAGGATATCCAGCTTCCCCGGTCTGGGGTGAGGAATGTGGGGAGC  
TTGGGCATCCTCCTCCAGCCTCCTCCAGCCCCAGGCAGTGCCTTACCTGTGGTGCCAGAAA  
AGTGCCCCCTAGGTTGGTGGGTCTACAGGAGCCTCAGCCAGGCAGCCACCCACCCCTGGGGCC  
CTGCCTCACCAAGGAAATAAAGACTCAAGCCATAAAAAAA

## **FIGURE 484**

MRPGAPGPLWPLPWGALAWAVGVSSMSGGNPAPGGVCWLQQGQEATCSLVLQTDVTRAECA  
SGNIDTAWSNLTHPGNKINLLGFLGLVHCLPCKDSCDGVCEGPGKACRMLGGRPRCECAPDCS  
GLPARLQVCGSDGATYRDECELRAARCRGHPDLVSVMYRGRCRKSCHEHVVCPRPQSCVVDQTGS  
AHCVVCRAAPCPVPSSPGQELCGNNNVTYISSCHMRQATCFLGRSIGVRHAGSCAGTPEEPPG  
GESAEEEENFV

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **N-glycosylation sites.**

amino acids 73-77, 215-219

#### **Osteonectin domain proteins.**

amino acids 97-130, 169-202



**FIGURE 485**

GCTCGAGGCCGGCGGGCGGGAGAGCGACCCGGGCGGCCTCGTAGCGGGGCCCCGGATCCCC  
GAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAG**ATG**ATGGGCTTGGGA  
AACGGGCGTCGCAGCATGAAGTCGCCGCCCTCGTGCTGGCCGCCCTGGTGGCCTGCATCATC  
GTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGATCATG  
GAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGGCGCCGTGGAGCTGAAGAAGAAC  
GAGTTCCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGCCACAAC  
TTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTTGGTGAATAACATC  
ACCACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGAGGAATTAC  
GGCAGGCTGCAGCAGGATGTCTCCAGTTTCAGAAGAACCAGACCAACCTGGAGAGGAAGTTC  
TCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTGTGAGGAGCGA  
ATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGAGTGAAAACAAC  
GACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAGGCAGCAGGCCTG  
CCACACACAGAGGTGCCACAAGGGAAGGGAAACGTGCTTGGTAACAGCAAGTCCCAGACACCA  
GCCCCCAGTTCCGAAGTGGTTTTTGGATTCAAAGAGACAAGTTGAGAAAGAGGAAACCAATGAG  
ATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAGCCAGGCCGGGAGCAG  
GTGGTGGAAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGGAGAACTGGGCCAGACC  
CCACAGGTGCAGGCTGCCCTGTCAGTGAGCCAGGAAAATCCAGAGATGGAGGGCCCTGAGCGA  
GACCAGCTTGTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCTGCCGGGGAAGGGAGAAAC  
CAGCAGAAACTGAGAGGAGAAGATGACTACAACATGGATGAAAATGAAGCAGAATCTGAGACA  
GACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAACATAGATGTTTTTAATGTTGAAGATCAG  
AAAAGAGACACCATAAATTTACTTGATCAGCGTGAAAAGCGGAATCATACACTC**TGA**ATTGAA  
CTGGAATCACATATTTCAACAACAGGGCCGAAGAGATGACTATAAAATGTTTCATGAGGGACTGA  
ATACTGAAAACGTGAAATGTACTAAATAAAATGTACATCTGA

**FIGURE 486**

MMGLGNGRRSMKSPPLVLAALVACIIIVLGFNYWIIASSRSVDLQTRIMELEGRVRRAAAERGAV  
ELKKNEFQGELEKQREQLDKIQSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQLKT  
LQRNYGRLQODVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVASRD  
LSENNDQRQQLQALSEPQPRLQAAGLPHTTEVPQGKGNVLGNSKSQTPAPSSEVVLDSCRQVEK  
EETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQENPEM  
EGPERDQLVIPDGQEEEQEAAGEGRNQQLRGEDDYNMDENEAESETDKQAALAGNDRNIDVF  
NVEDQKRDTINLLDQREKRNHTL

**Important features:**

**Signal peptide:**

amino acids 1-29

**FIGURE 487**

AACTCAAACCTCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGGG  
TGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATG**TAT  
GGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGGCCT  
ATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGCTCGG  
TTAAAATGCACTTTCTCCAGCTTTGCCCTGTGGGTGATGCTCTAACAGTGACCTGGAATTTT  
CGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTCCAACCC  
ATGAGTGGGCGGTTTAAGGACCGGTGTCTTGGGATGGGAATCCTGAGCGGTACGATGCCTCC  
ATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGAAGAACCCA  
CCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTACGCTTCTCT  
GAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCATAATAGTAATT  
GTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTCATAAAGTGGTG  
GAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCTGTTTATTTAGAA  
GACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAAGAACCCTAGTATT  
TCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGTTTTTCCAACCAGTTC  
TGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCACAGTGCTCCTCCATAT  
CACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTTCAAGAGTGTAATTTTTTCAA  
GTGCTCATTAGGTTTTTATAAACAAGAAGCTACATTTTTTGCCCTTAAGACACTACTTACAGTGT  
TATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCCAATTTGTCTGTACATTT  
CCTTTCACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTAATGTGTTTACTCTCTTTCC  
TTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTGTTTCTGATTAACAGTAAATC  
CTAAATTCAAACCTGTTAAATGACATTTTTATTTTTATGTCTCTCCTTAACCTATGAGACACATC  
TTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATTTTTTGTCG

488/550

## **FIGURE 488**

MYGKSSTRAVLLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVTW  
NFRPLDGGPEQFVFYYHIDPFQPMSEGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQVK  
NPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLFFQHYRKKRWAERAHK  
VVEIKSKEEERLNQEKKVSVYLETD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 489**

[illegible]

**FIGURE 490**

MLLLWVSVVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDSFDGRLTFHPGSQVVKLPFINFM  
KTRGTSFLNAYTNSPICCPSRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQKF  
GKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNWLRK  
EAINYTEPFVIYLG LNLPHYPSPSSGENFGSSTFHTSLYWLEKVSHDAIKIPKWSPLSEMHP  
VDYYSSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQDLLQKTIVIIYSSDHG  
ELAMEHRQFYKMSMYEASAHVPLLMMGPGIKAGLQVSNVSLVDIYPTMLDIAGIPLPQNLSG  
YSLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYS DGASILPQLFD  
LSSDPDEL TNVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIGQNYSNVIA  
NLRWHQDWQKEPRKYENAI DQWLKTHMNPRAV

**Important features:****Signal peptide:**

amino acids 1-15

**N-glycosylation sites.**

amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,  
498-501

**Sulfatases proteins:**

amino acids 286-315, 359-369, 78-97

**FIGURE 491**

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGA  
GCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGC  
CTCTCTTGGCCTCCAACCTTGTTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGT  
TGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTTGTGACAGCAGT  
TGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGA  
CATCTATAGCACCCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCCAGGCCATGATGGTGAC  
ATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTT  
CTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGG  
AGGCCCTCCTGGGATTCAATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTC  
ACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTT  
TTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCTGCTCATCCCAGAGAAATCG  
CTCCAACCTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGG  
TCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCCTACAGCCTGACAGGGTATGTGTGAAGAAC  
CAGGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAGCACCCCCGAGGGCCA  
CAGGTGAGGGACACTACCACTGGATCGTGTGAGAAGGTGCTGCTGAGGATAGACTGACTTTTG  
CCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATGCAGGTTGAATTGCCAA  
GGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTCCCCTGCCCTAAGTCCCC  
AACCCTCAACTTGAAACCCCATTCCTTAAAGCCAGGACTCAGAGGATCCCTTTGCCCTCTGGT  
TTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCCTGACTGACCCCTCTGTGATCAA  
AGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGGGGATGGGAAGGAGAAGCAGT  
GGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCCAAAGAACTGATTGGCCC  
TGGAACCTCCATCCCCTCTTGTTATGACTCCACAGTGTCCAGACTAATTTGTGCATGAACTG  
AAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATGCAGGATGGGAGGACAGGAA  
GGCAGCCTGGGACATTTAAAAAAATA

## **FIGURE 492**

MASLGLQLVGYYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQ  
CDIYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFI  
LGGLLGFIPIVAVNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQR  
NRSNYYDAYQAQPLATRSSPRPGQPPKVKSEFNSSYSLTGYV

### **Important features:**

#### **Signal peptide:**

amino acids 1-24

#### **Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

#### **N-glycosylation site.**

amino acids 190-193

#### **PMP-22 / EMP / MP20 family proteins.**

amino acids 46-59



**FIGURE 493**

GCACTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCCATGGTGCCCAGAATCTTCGCTCCTGCT  
TATGTGTCAGTCTGTCTCCTCCTCTTGTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAGAA  
CCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAGTGC  
TGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCCTGCACCTTC  
TGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCCCTTTGGCCTCACAAACGATTTTGTGTGAAG  
CTGAAGGTTTCAGGGTGTGAATTCCCAGTGCCACTCATCTCCCATCTCCAGTAAATGTGAAAGC  
AGAAGACGTTTTCCCTTGAGAAGACATAGAAAGAAAATCAACTTTCACTAAGGCATCTCAGAAA  
CATAGGCTAAGGTAATATGTGTACCAGTAGAGAAGCCTGAGGAATTTACAAAATGATGCAGCT  
CCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGTAGATTATCA  
GGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

**FIGURE 494**

MVPRI FAPAYVSVCLLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYND AIVSLSET  
RQCGPPCTFWPCFELCCLDSFGLTND FVVKLKVQGVNSQCHSSPISSKCESRRRFP

**Important features:**

**Signal peptide:**

amino acids 1-25

**FIGURE 495**

CTCCACTGCAACCACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATTT  
TCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCTGA  
TGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCTGCAGCACTGTTGCT  
ATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAACTGCACCTTCAGAGTCT  
GCTTTGAGCAGTGCTGCCCCCTGGACCTTCATGGTGAAGCTGATAAACCAAGAACTGCGACTCAG  
CCCGGACCTCGGATGACAGGCTTTGTGCGAGTGTGAGCTAATGGAACATCAGGGGAACGATGA  
CTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTACCTGAGATCTGGGA  
TGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCCTTCATTCTGTGACC  
TGTCTGAGGCCACCCTGCAGCTGCCCTGAGGAGGCCACAGGTCCCCTTCTAGAATTCTGGA  
CAGCATGAGATGCGTGTGCTGATGGGGGCCAGGGACTCTGAACCCTCCTGATGACCCCTATG  
GCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCTTCACCCTTCTGTGAGATT  
TTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAATAAATTTATGTA  
CTTTATAAATGAAAA

TGGTAT = 248200T

496/550

## **FIGURE 496**

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVPL  
ARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCDSARTSDDRLLCRSVS

**Important features:**

**Signal peptide:**

amino acids 1-24

496/550

**FIGURE 497**

[illegible]

**FIGURE 498**

MKVVPSLLLSVLLAQVWLVPGLAPSPQSPETPAPQNQTSRVVQAPREEEDEQEASEEEKAGEE  
EKAWLMASRQQLAKETSNFGFSLLRKISMRHDGNMVFSFPGMSLAMTGLMLGATGPTETQIKR  
GLHLQALKPTKPGLLPSLFKGLRETLNRNLELGLSQGSFAFIHKDFDVKETFFNLSKRYFDTE  
CVPMNFRNASQAKRLMNHYINKETRGKIPKLFDEINPETKLILVDYILFKGKWLTDFDPVFTE  
VDTFHLDKYKTIKVPMMYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHLALEDY  
LTTDLVETWLRNMKTRNMEVFFPKFKLDQKYEMHELLRQMGIRRIFFSPFADLSELSATGRNLQ  
VSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLFLGRVVNP  
TLL

**FIGURE 499**

CTAGCCTGCGCCAAGGGGTAGTGAGACCGCGCGGCAACAGCTTGCGGCTGCGGGGAGCTCCCG  
TGGGCGCTCCGCTGGCTGTGCAGGCGGCC**ATG**GATTCCTTGCGGAAAATGCTGATCTCAGTCG  
CAATGCTGGGCGCAGGGGCTGGCGTGGGCTACGCGCTCCTCGTTATCGTGACCCCGGAGAGC  
GGCGGAAGCAGGAAATGCTAAAGGAGATGCCACTGCAGGACCCAAGGAGCAGGGAGGAGGCGG  
CCAGGACCCAGCAGCTATTGCTGGCCACTCTGCAGGAGGCAGCGACCACGCAGGAGAACGTGG  
CCTGGAGGAAGAACTGGATGGTTGGCGGCGAAGGCGGCGCCAGCGGGAGGTCACCG**TG**AGACC  
GGACTTGCCCTCCGTGGGCGCCGGACCTTGGCTTGGGCGCAGGAATCCGAGGCAGCCTTTCTCC  
TTCGTGGGCCCAGCGGAGAGTCCGGACCGAGATACCATGCCAGGACTCTCCGGGGTCCTGTGA  
GCTGCCGTCGGGTGAGCACGTTTCCCCCAAACCCTGGACTGACTGCTTTAAGGTCCGCAAGGC  
GGGCCAGGGCCGAGACGCGAGTCGGATGTGGTGAAC TGAAAGAACCAATAAAATCATGTTTCCT  
CCAAA

500/550

## **FIGURE 500**

MDSLRKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLAT  
LQEAATTQENVAWRKNWMVGGEGGASGRSP

**Important features:**

**Signal peptide:**

amino acids 1-18

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.





**FIGURE 502**

MGPSTPLLILFLLSWGPLQGQHHLVEYMERRLAALAEERLAQCQDQSSRHAAELRDFKNKML  
PLLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGRRN  
EKYDMVTDCGYTISQVRSMKILKRFGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRLRDF  
TLAMAARKASRVRVFPWPVGTGQLVYGGFLYFARRPPGRPGGGGEMENTLQLIKFHLANRTVV  
DSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCLAKLDPQTLDTQQWDTP  
CPRENAEAAFVICGTLYVVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAHASLRYN  
PRERQLYAWDDGYQIVYKLEMRKKEEV

**Important features:**

**Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 177-180, 248-251

**FIGURE 503**

TGCGGCGCAGTGTAGACCTGGGAGG**ATG**GGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGGT  
CTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGGCC  
CTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAAGAA  
CGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCCTCTCAGCA  
CGGGCTGGGAGGGTGTGACCAGAGTGTGATGGACCTGATAAAGCGAAACTCCGGATGGGTGTT  
TGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGACTATGC  
CATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGTACAGTCT  
GACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTCACCAAGTGGAGCAGGAGCCTGGGCTT  
CCTGTCACAG**TAG**CAGGCCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAGATCCTTCTGTG  
AGTGCTGCGTCCCCAGTAGGGATGGCGCCACAGGGTCCTGTGACCTCGGCCAGTGTCCACCC  
ACCTCGCTCAGCGGCTCCCGGGGCCAGCACCAGCTCAGAATAAAGCGATTCCACAGCA

## **FIGURE 504**

MGGLLLAAFLALVSVPRQAQAVWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVTL  
TPENNLRTLSSQHGLGGCDQSVMDLIKRNSGWVFENPSIGVLELWVLATNFRDYAIIIFTQLEF  
GDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

**Important features:**

**Signal peptide:**

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 505**

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCACA  
GCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCC**ATG**AGGATTCTGCAGTTAATC  
CTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAGTGC  
AAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTCGAGAAGACGCGGCTACTCTGTGGGGCG  
ACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACATAGTT  
CACCTGGGGCAGCACAACCTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCCACTGAG  
TCCTTCCCCCACCCGGCTTCAACAACAGCCTCCCCAACAAGACCACCGCAATGACATCATG  
CTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCTCTCCTCA  
CGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCCAGCCCCCAG  
TTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTGAGCACCAGAAGTGTGAG  
AACGCCTACCCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGGAAGGGGGCAAG  
GACTCCTGCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCAGTCTCTTCAAGGCATTATC  
TCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACACGAAAGTCTGCAAA  
TATGTGGACTGGATCCAGGAGACGATGAAGAACAAT**TAG**ACTGGACCCACCCACCACAGCCCA  
TCACCCTCCATTTCCACTTGGTGTTCCTGTTCACTCTGTTAATAAGAAACCCTAAGCC  
AAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATGCTGTCACTTAATAATC  
AACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTTGAAATATTGTGACTCTG  
GGAATGACAACACCTGGTTTGTTCCTCTGTTGTATCCCCAGCCCCAAAGACAGCTCCTGGCCAT  
ATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAA

## **FIGURE 506**

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAHC  
LKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPV SITWAV  
RPLTLSSRCV TAGTSC LISGWGSTSSPQLRLPHTLR CANITII EHQKCENAYPGNITDTMVCA  
SVQEGGKDSCQGD SGGPLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDW IQETMKNN

### **Important features:**

#### **Signal peptide:**

amino acids 1-18

#### **Serine proteases, trypsin family, histidine active site.**

amino acids 58-63

#### **N-glycosylation sites.**

amino acids 99-102, 165-168, 181-184, 210-213

#### **Glycosaminoglycan attachment site.**

amino acids 145-148

#### **Kringle domain proteins.**

amino acids 197-209, 47-64

#### **Serine proteases, trypsin family, histidine protein**

amino acids 199-209, 47-63, 220-243

#### **Apple domain proteins**

amino acids 222-249, 189-222









510/550

**FIGURE 510**

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPDHT  
NETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQNTS  
QISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSILYIGCKM  
YYSRRGIRYRTIDEHDAII

10023072 131501



**FIGURE 512**

MARMSFVIAACQLVLGLLMTSLTESSIONSECPQLCVCEIRPWFTPQSTYREATTVDCNDLRL  
 TRIPSNLSSDTQVLLLQSNNIAKTVDELQQLFNLTELDIFSQNNFTNIKEVGLANLTQLTTLHL  
 EENQITEMTDYCLQDLSNLQELYINHNQISTISAHAFAGLKNLLRLHLNSNKLKVIDSRWFDS  
 TPNLEILMIGENPVIGILDMNFKPLANLRSLVLAGMYLTDIPGNALVGLDSLESLSFYDNKLV  
 KVPQLALQKVPNLKFLDLNKNPIHKIQEGDFKNMLRLKELGINNMGELVSVDRYALDNLPELT  
 KLEATNNPKLSYIHRLAFRSVPALESLMLNNNALNAIYQKTVESLPNLREISIHNSNPLRCDCV  
 IHWINSNKTNIRFMEPLSMFCAMPPEYKGHQVKEVLIQDSSEQCLPMISHDSFPNRLNVDIGT  
 TVFLDCRAMAEPEPEIYWVTPIGNKITVETLSDKYKLSSEGTLEISNIQIEDSGRYTCVAQNV  
 QGADTRVATIKVNGTLLDGTQVLKIYVKQTESHSILVSWKVNSNVMTSNLKWSSATMKIDNPH  
 ITYTARVPVDVHEYNLTHLQPSTDYEVCLTVSNIHQQTQKSCVNVTTKNAAFVAVDISDQETST  
 ALAAVMGSMFAVISLASIAVYFAKRFRKKNYHHSLLKKYMQKTSSIPLNELYPPLINLWEGDSE  
 KDKDGSADTKPTQVDTSRSYYMW

**Important features:****Signal peptide:**

Amino acids 1-25

**Transmembrane domain:**

Amino acids 508-530

**N-glycosylation sites:**

Amino acids 69-73;96-100;106-110;117-121;385-389;517-521;  
 582-586;611-615

**Tyrosine kinase phosphorylation site:**

Amino acids 573-582

**N-myristoylation sites:**

Amino acids 16-22;224-230;464-470;637-643;698-704

**FIGURE 513**

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTCC  
CAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCCCT  
GTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAACCGAGGGCT  
GAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGGAATCACGCATGC  
CGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGCAAGGA  
GTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGATCAACCA  
TGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACGCTGCTGG  
ACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCACCAGGCTGG  
GAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGGAAAGGAAGT  
GGAGAAGCTTGGCCAAGGTGCCCACCATGCTGCTGGCCAGGCCGGGAAGGAGCTGCAGAATGC  
TCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGCAACCATCAAAG  
CGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGGGGCCTCAGTCAA  
CACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCATGCCCC**TAA**ACTGG  
CATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTACATCAGCTGACATGACCTGGAGGGG  
TTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTTGTACTGGGATTTGTGAATAAA  
CTTGATACACCA





## **FIGURE 516**

MPSSTAMAVGALSSSLLVTCCLMVALCSPSIPLEKLAQAPEQPGQEKREHATRDGPGRVNELG  
RPARDEGGSGRDWKS KSGRGLAGREPWSKLQAWVSQGGGAKAGDLQVRPRGDT PQAEALAAA  
AQDAIGPELAPTPEPP EYVYPDYRGKGCVD ESGFVYAIG EKFA PGPSAC PCLCTEEGPLCAQ  
PECPRLHPRCIHVDT SQCCPQCKERKNYCEFRGKTYQTLEEFVVSPCERCRC EANGEVLCTVS  
ACPQTECVDPVYEPDQCCPICKNGPNCFAETAVIPAGREVKTD ECTICHCTYE EGTWRIERQA  
MCTRHECRQM

### **Important features:**

#### **Signal peptide:**

amino acids 1-27

#### **Transmembrane domain:**

amino acids 11-30

#### **Glycosaminoglycan attachment site.**

amino acids 80-83

#### **N-myristoylation sites.**

amino acids 10-15, 102-107, 103-108

#### **Cell attachment sequence.**

amino acids 114-117

#### **EGF-like domain cysteine pattern signature.**

amino acids 176-187



**FIGURE 517**

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCCT  
TTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCTAC  
GGACGACGCCT**ATGA**AAGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTGTGC  
CGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAACTTGAATCATTATATACAAG  
TTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAAATCTAACT  
CTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAATTTAAGGAGCTAGTTACACATG  
GAGACGCTTCAACTGAGAATGATGTTTTTAACCAATCCTATCAGTGAAGAACTACAACCTTCC  
CTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAACACACGGAAAGTACCCCATTTCTGGT  
CGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTGAAAATGAAG  
AGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAACTGAGGCACCAAGAATGTTGCCAGTTG  
TACTGAATCATCTACAAGTCCATATGTTACCTCATACAAGTCACCTGTCACCACTTTAGATA  
AGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCCTCAGCTCTCAGGTGAACTG  
CGATAGAAAAACCCGAAGAGTTTGGAAAGCACCCAGAGAGTTGGAATAATGATGACATTTTGA  
AAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACCTTCTTAGTGACACCAGCAACCCAG  
CATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCCTTGCTCTAGCAGCAG  
CAGCAGAACATAAATTAAAAACAATGTATAAGTCCCAGTTATTGCCAGTAGGACGAACAAGTA  
ATAAAATTGATGACATCGAAACTGTTATTAACATGCTGTGTAATTCTAGATCTAACTCTATG  
AATATTTAGATATTAAATGTGTTCCACCAGAGATGAGAGAAAAAGCTGCTACAGTATTCAATA  
CATTAAAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTAAAAGTTTAT**TAA**ACAATAA  
TATAAAAATTTTAAACCTACTTGATATTCCATAACAAAGCTGATTTAAGCAAACCTGCATTTTT  
TCACAGGAGAAATAATCATATTCGTAATTTCAAAGTTGTATAAAAATATTTTCTATTGTAGT  
TCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACAATTTTCATATGCACTAAAAACC  
TAATTTAAAATAAAATTTTGGTTCAGGAAAAA

## **FIGURE 518**

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRSVPSGEPGREKKSNSPKH  
VYSIASKGSKFKELVTHGDASTENDVLTNPISEETTTFTPTGGFTPEIGKKKHTESTPFWSIKP  
NNVSIVLHAEOPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSTSPYVTSYKSPVTTLDKSTG  
IEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNPAYRE  
DIEASKDHLKRSLALAAAAEHKLKTMYSQLLPVGRTSNKIDDIETVINMLCNSRSKLYEYLD  
IKCVPPERMREKAATVFNTLKNMCRSRRVTALLKVY

**Important features:**

**Signal peptide:**

amino acids 1-19





**FIGURE 521**

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATGG  
TTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGGAA  
ACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CAGGA  
TGAAGATGGATACATCACCTTAAATATTTAAAACTCGGAAACCAGCTCTCGTCTCCGTTGGCCC  
TGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGGATGGT  
TGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGATGAGAA  
TGAAAATCGCACAGGAACTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGGTAAAACA  
ATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAACTGGAGATA  
TTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAGTAAGCAGTA  
CTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGGAGTACATCAA  
AGCCAGGACTCATTTAATTCGTTGGGTCTGGATTATCTCGCCAGAAGTCGAATGAGGTCTGGAA  
GTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGGAAGATGGAAAAGGAAA  
TATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTGAGAACAAACATTA  
TTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT**TAA**TGCAAAGAGGT  
GGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGTATGAATGCATCAGTA  
GCTGAAAAAAAAAAAAA



**FIGURE 523**

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATGGC**  
AAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAA  
GAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAAT  
TGTCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGA  
GCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAG  
AACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAA  
CGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGT  
GATTCCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTT  
CTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAA  
AAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAAT  
ATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGA  
AAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCG  
TCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTGAAAATGGAATAGA  
ATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTA  
TTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATACTGCTACCAAGGAGG  
ACGAGTCATCTGTCGTGTCATCATGCCTTGTAAGTGGTGGGTGGCCCGCATGCTGGGGAGGGT  
**CTAA**TAGGAGGTTTGAGCTCAAATGCTTAACTGCTGGCAACATATAATAAATGCATGCTATT  
CAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCTCTCCAGAATTACTTGTAG  
GTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACCAAAAAAAAAAAAAAAAAA

**FIGURE 524**

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLLIVLFWGSKHFWPEVPPKAYD  
 MEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQI  
 KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPT  
 LISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENG  
 IEFDPMLDERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

**Important features:****Signal peptide:**

amino acids 1-40

**Transmembrane domain:**

amino acids 25-47 (type II)

**N-glycosylation sites.**

amino acids 94-97, 180-183

**Glycosaminoglycan attachment sites.**

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

**N-myristoylation sites.**

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

**Microbodies C-terminal targeting signal.**

amino acids 315-317

**Cytochrome c family heme-binding site signature.**

amino acids 9-14



**FIGURE 525**

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGCATGAAGATCACTGGGGGTC  
TCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCCAA  
AAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCACAT  
ACCTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACTTGTGTACCGAGA  
GCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCTAAATTCTCCATGGAC  
ATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTGAGTTTC  
TTTCAGTTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTGACTGAAT  
GGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCATTTTTTTTTT  
TAACACGTCAATAAAAAAATAATCTCCAGA





**FIGURE 528**

MLRAPGCLLRTSVAPAAALAAALLSSLARCSLLEPRDPVASSLSFYFGTKTRYEDVNPVLLSG  
PEAPWRDPELLEGTCTPVQLVALIRHGTRYPTVKQIRKLRQLHGLLQARGSRDGGASSTGSRD  
LGAALADWPLWYADWMDGQLVEKGRQDMRQLALRLASLFPALFSRENYGRLRLITSSKHRCMD  
SSAAFLQGLWQHYPGLPPPDVADMEFGPPTVNDKLMRFFDHCEKFLTEVEKNATALYHVEAF  
KTGPEMQNILKKVAATLQVPVNDLNADLIQVAFFTCSEFDLAIKGVKSPWCDVFDIDDAKVLEY  
LNDLKQYWKRGYGYTINSRSSCTLFQDIFQHLDKAVEQKQRSQPISSPVILQFGHAETLLPLL  
SLMGYFKDKEPLTAYNYKKQMHRKFRSGLIVPYASNLI FVLYHCENAKTPKEQFRVQMLLNEK  
VLPLAYSQETVSFYEDLKNHYKDILQSCQTSEECELARANSTSDDEL

**Important features:****Signal sequence**

amino acids 1-30

**N-glycosylation sites.**

amino acids 242-246, 481-485

**N-myristoylation sites.**

amino acids 107-113, 113-119, 117-123, 118-124, 128-134

**Endoplasmic reticulum targeting sequence.**

amino acids 484-489

**FIGURE 529**

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGGTGG  
 CGGAGAGATCAGAAGCCTCTTCCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCTCAGGGA  
 CGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACG**ATG**GCAGCGTCCGCCGGAGCCGGGGCG  
 GTGATTGCAGCCCCAGACAGCCGGCGCTGGCTGTGGTGGTGGTGGCGGGCGGCGCTTGGGCTC  
 TTGACAGCTGGAGTATCAGCCTTGAAGTATATACGCCAAAAGAAATCTTCGTGGCAAATGGT  
 ACACAAGGGAAGCTGACCTGCAAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACCTCAGTC  
 TCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTCGTTTTTCCACTACTCCCAAGGG  
 CAAGTGTACCTTGGGAATTATCCACCATTAAAGACAGAATCAGCTGGGCTGGAGACCTTGAC  
 AAGAAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACACAATGGCACCTATATCTGT  
 GATGTCAAAAACCTCCTGACATCGTTGTCCAGCCTGGACACATTAGGCTCTATGTCGTAGAA  
 AAAGAGAATTTGCTGTGTTTCCAGTTTGGGTAGTGGTGGGCATAGTTACTGCTGTGGTCTTA  
 GGTCTCACTCTGCTCATCAGCATGATTCTGGCTGTCCTCTATAGAAGGAAAACTCTAAACGG  
 GATTACACTGGCTGCAGTACATCAGAGAGTTTGTACCAGTTAAGCAGGCTCCTCGGAAGTCC  
 CCTCCGACACTGAGGGTCTTGTAAGAGTCTGCCTTCTGGATCTCACCAGGGCCCAGTCATA  
 TATGCACAGTTAGACCACTCCGGCGGACATCACAGTGACAAGATTAACAAGTCAGAGTCTGTG  
 GTGTATGCGGATATCCGAAAGAAT**TAA**GAGAATACCTAGAACATATCCTCAGCAAGAAACAAA  
 ACCAAACTGGACTCTCGTGCAGAAAATGTAGCCCATTACCACATGTAGCCTTGGAGACCCAGG  
 CAAGGACAAGTACACGTGTACTCACAGAGGGAGAGAAAGATGTGTACAAAGGATATGTATAAA  
 TATTCTATTTAGTCATCCTGATATGAGGAGCCAGTGTTGCATGATGAAAAGATGGTATGATTC  
 TACATATGTACCCATTGTCTTGCTGTTTTTGTACTTTCTTTTCAGGTCATTTACAATTGGGAG  
 ATTTTCAGAAACATTCCTTTTACCATCATTTAGAAATGGTTTGCCTTAATGGAGACAATAGCAG  
 ATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAACTAAGGGCTTAAGACTGATTAGTCTTA  
 GCATTTACTGTAGTTGGAGGATGGAGATGCTATGATGGAAGCATACCCAGGGTGGCCTTTAGC  
 ACAGTATCAGTACCATTTATTTGTCTGCCGCTTTTAAAAAATACCCATTGGCTATGCCACTTG  
 AAAACAATTTGAGAAGTTTTTTTTGAAGTTTTTCTCACTAAAATATGGGGCAATTGTTAGCCTT  
 ACATGTTGTGTAGACTTACTTTAAGTTTTGCACCCTTGAAATGTGTCATATCAATTTCTGGATT  
 CATAATAGCAAGATTAGCAAAGGATAAATGCCGAAGGTCACCTCATTCTGGACACAGTTGGAT  
 CAATACTGATTAAGTAGAAAATCCAAGCTTTGCTTGAGAACTTTTGTAACGTGGAGAGTAAAA  
 AGTATCGGTTTTTA

## **FIGURE 530**

MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKST  
STTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKDDASINIENMQ  
FIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLTLLISMILAV  
LYRRKNSKRDTGCGSTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSGGHHS  
DKINKSESVVYADIRKN

**Important features:**

**Signal peptide:**

amino acids 1-37

**Transmembrane domain:**

amino acids 161-183

1003072-131901

**FIGURE 531**

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCTCGA  
GGCTGGTGGGAAGAAGCCGAGATGGCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGCTC  
TTGCTGATGGCGGTAGCAGCGCCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACTGGT  
GCGCGAGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCTGTGGCACGGTGGGGCTGCTGCTG  
GAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCTGGAAC  
CAGCAGGATGGTACCTTGTCCCTGTCACAGCGGCAGCTCAGCGAGGAGGAGCGGGGCCGACTC  
CGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGGGGCCCTG  
GATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGGTGGAGTCG  
CACCTGTCTGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGCGTGTCTGGTG  
GTGACGCACCCCGGGGGCTGCCGGGGCCATGAGGTGGAGGACGTGGACCTGGAGCTGTTCAAC  
ACCTCGGTGCAGCTGCAGCCGCCACCACAGCCCCAGGCCCTGAGACGGCGGCCTTCATTGAG  
CGCCTGGAGATGGAACAGGCCCAGAAGGCCAAGAACCCCCAGGAGCAGAAGTCCTTCTTCGCC  
AAATACTGGATGTACATCATTCCTCGTCCTGTTCCCTCATGATGTCAGGAGCGCCAGACACC  
GGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCCTTTGCTGTGTGCCA  
CCCTCCCTGTAAGTCTATTTAAAACATCGACGATACATTGAAATGTGTGAACGTTTTGAAAA  
GCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGGTCCTGATGTACAAGCT  
TGATTGAAATTCAGTCTCACTTGATACGTTATTCAGAAACCCAAGGAATGGCTGTCCCCATC  
CTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTATTAAACTGTCCCCCAGATC  
GACACGCAAAAAAAAAA







## **FIGURE 534**

MELALLCGLVVMAGVPIQGGILNLNKMVKQVTGKMPILSYWPGCHCGLGGRGQPKDATDWC  
CQTHDCCYDHLKTQGCGLIYKDNNKSSIHCMDL SQRYCLMAVFNVIIYLENEDSE

**Important features:**

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 1-24

**N-glycosylation site.**

amino acids 86-89

**N-myristoylation sites.**

amino acids 20-25, 45-50

**Phospholipase A2 histidine active site.**

amino acids 63-70

**FIGURE 535**

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTG  
 AACTGGGTGCTCATCACGGGAAGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCC  
 CCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTT  
 TTTTAAACGCCCCCTCCCCACCCCCCAAAAACTGTAAAGATGCAAAAACGTAATATCCAT  
 GAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTT  
 GTTCTTGGAGTGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAG  
 GGGTCCAATTTTTCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGG  
 CTGTCATGCAACTGGCCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAA  
 AGG**ATG**GGTTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCC  
 ACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAA  
 GGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGT  
 TGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGG  
 CTAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATATTGACGAAAATGCT  
 TTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTT  
 AACAATACCTTCAGACCTGTGACAAATTTACGGAACCTGGATCTGTCTTATAATCAGCTGCAT  
 TCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTTACATTTACGGTCTAAC  
 TCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACCTGGAACCTTTTGACCTG  
 GGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGCATGATCAGACTCAAAGAA  
 CTTACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCTTTTCCAAGGTTGGTCAGC  
 CTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTATAGGACAGACCATGTCTGGACC  
 TGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATCGAAGCTTTCAGTGGACCCAGT  
 GTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGATTCCAACAAGCTCACATTTATT  
 GGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAGTCTTGCTGGGAATATATGG  
 GAATGCAGCAGAAATATTTGCTCCCTTGTAAACTGGCTGAAAAGTTTTAAAGGTCTAAGGGAG  
 AATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTGATCGATGCAGTGAAG  
 AACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAGGGCTCTCCCAAAG  
 CCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTTTGCCCCCGACG  
 GTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCTTTCCATAAA  
 ATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTATCTACGTG  
 TCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAAGGCAC  
 AGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTAGAT  
 TATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTAT  
 AACAAATCGGGCTCCAGGGAGTGTGAGGTAT**TGA**ACCATTGTGATAAAAAGAGCTCTTAAAGC  
 TGGGAAATAAGTGGTGCTTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTC  
 CCCTTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTTGTCCGTTTTAGTGCATTATA  
 ATACTGGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGT  
 GAAGCTTGAACCTCGGTTTAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAAT  
 GTCGCATTTGTTTTAAGATAAAACTTCTTTCATAGGTAAAAA



**FIGURE 537**

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCCTCAGCAACCCTCGAC**ATG**GCGCTGAGGCGGCCACCGCGAC  
 TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCTGCTGCTGCTTTTTAGGGGGCTGCCTGATAGGGGGCTGTAAATC  
 TCAAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGTGGAACCTGTCTTGCATCATTACGGATTTCGC  
 AGACAAGTGACCCCAGGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTGTACAAACAAA  
 TTCAGGGAGACTTGGCGGGTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG  
 ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA  
 CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC  
 ACTGCCAGGAGAGTGAGGGCCACCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCACCGGATT  
 CCAGAGCCAATCCAGATTTTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCACTGCTG  
 TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG  
 AGATGGAAGTCTATGACCTGAACATTGGCGGAATTATTGGGGGGGTTCTGGTTGTCTTGTCTGACTGGCCCTGA  
 TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA  
 ACCCAGGGAAACCAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTTG  
 TGATCT**TGA**GACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA  
 GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTGAGAAGCTTTTCGTTTTTGGCCAAAGTTGACCA  
 CTACTCTTCTTACTCTAACAAGCCACATGAATAGAAGAATTTTCTCAAGATGGACCCGGTAAATATAACCACAA  
 GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTTCTAATCTGTTTTCTGGCCTGATTCCCGCATGAGTATTAGG  
 GTGATCTTAAAGAGTTTGCTCACGTAAACGCCCCGTGCTGGGCCCTGTGAAGCCAGCATGTTCCACCACTGGTCGTT  
 CAGCAGCCACGACAGCACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGGCGGGAACCCA  
 GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCACAGACACCACCGCAGTTTTCTTCTTAAAGGCTCTGC  
 TGATCGGTGTTGCAGTGTCATTGTGGAGAAGCTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG  
 GTAAATTGGTTGCTGGAAGAGGGATCTTGCTGAGGAACCTGTCTGTCCAACAGGGTGTGAGGATTTAAGGAAA  
 ACCTTCGTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTTCTATGGGTCTTGTTTTATTTTATAAAATTT  
 TACATCTAAATTTTTTGCTAAGGATGTATTTTGATTATTGAAAAGAAAATTTCTATTTAACTGTAAATATATTGT  
 CATACAATGTTAAATAACCTATTTTTTTTAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTTAAAT  
 TGGAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT  
 CACACAAGTTTTTAGCCTTTTTTACAAGGGAACCTACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT  
 TAAAAATTCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT  
 GCCTCTTCCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG  
 CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCGCTCTAGCTCACTGTTGCCCTCGCTGTCTGCCAGGAGGCCCT  
 GCCATCCTTGGGCCCCGGCAGTGGCTGTGTCCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT  
 TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT  
 TTTTGGTTATGGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTTAAAGTTTGTTTAATTATTTGTT  
 AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTTAAAGAAAATGGAT  
 CCCACTGTTTCCTCTTTGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTGCGATTTCAAAACAAACCATGAT  
 GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTCAGGTGGAGCAGCCAGGTGAAAGGCCCTGGCGGGGAGGAAAG  
 TGAAACGCCTGAATCAAAAGCAGTTTTTCTAATTTTGACTTTAAATTTTTTCATCCGCCGGAGACACTGCTCCCATT  
 TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT  
 GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCTGGA  
 GAATGGCTCTCACTACTCACCTTGTCTTTAGCTTCCAGTGCTTGGGTTTTTTTATACTTTGACAGCTTTTTTTT  
 AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTTATGTGAAACACTTTGCCGCAGGCCGCTGGCAGAGGCA  
 GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCATGGCATCCTGGATGCTTAGCATGCAAGTTC  
 CCTCCATCATTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTACGCTCCAGCCTCCT  
 TCTTGGTTGTATAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTATACCCTAAAACTTCTACACTAGTGCCA  
 TGGGAACCAGGTCTGAAAAAGTAGAGAGAAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA  
 CGGAAAAGGAATACTCGTGATTTTAAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT  
 GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTTGTTTAAAC  
 CTCATTTATAAAAGCTTCAAAAAACCCA

**FIGURE 538**

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTSDF  
RIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLSKIWNVTRRDSALYRCEVVARNDRKEI  
DEIVIELTVQVKPVTVPVCRVPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRANPR  
FRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVYDLNIGGIIGGVLVV  
LAVLALITLGICCAYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

**Important features:****Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 243-263

**N-glycosylation sites.**

amino acids 104-107, 192-195

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 107-110

**Casein kinase II phosphorylation site.**

amino acids 106-109, 296-299

**Tyrosine kinase phosphorylation site.**

amino acids 69-77

**N-myristoylation sites.**

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267

**FIGURE 539**

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTAGAGGCCGGGGAAGAGAAGCAAAGC  
 GCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCCCTAA  
 CTTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAACCTCCAGCCCCGCACATCCACGCGCGG  
 CACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCTGGGCAGCTGG  
 GCTCGGGCGGGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGG  
 CTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGA  
**TG**AGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCC  
 GCCGCGTGGTCTAGCGGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGG  
 CCTACTTCCATGAACTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTG  
 AGGGAGGAGTCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAAGTTAATAGAGAGCATGTTGC  
 AAAACCTGACAAAACCCGGGACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGA  
 ATGGAGATGGGCAAACATCTGGTGCCTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATT  
 CCCAGTACCGAAACTGGTACACAGATGAACCTTCTGCGGAAGTGAAAAGTGTGTTGTGATGT  
 ATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAAATGATGACA  
 GGTGTAACATGAAGCACAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCCTG  
 TAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAG  
 CAGGTATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATAC  
 TGGTTGCTTTTGGAACTGTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAAACTA  
 GTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGTAT  
**AA**ATAACTCATTGACTTGGTTCAGAAATTTTGTAAATCTGGATCTGTATAAGGAATGGCATCAG  
 AACAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGCAAGATGAACTGTAAGCTCCCCCT  
 TGAGGCAAATATTAAAGTAATTTTTATATGTCTATTATTTTCAATTTAAAGAATATGCTGTGCTA  
 ATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAAACTTCAAACCTTCAAGCAAA  
 TGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTGCGGAGTATGTGTGTTAGAAGCAAT  
 TCCTTTTATTTCTTTACCTTTTCATAAGTTGTTATCTAGTCAATGTAATGTATATTGTATTGA  
 AATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAAATGAACTGTTCTA  
 ATATTTATTTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAGAACTTATTAC  
 TGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGTCTCTCGAA  
 ATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGAAATAAGA  
 AGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGATTGTCT  
 AATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTTAGCTTAAAATTAACAGATTTTG  
 TAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAGT  
 GACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTG  
 AGGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCTGGGGTTT  
 GGGATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGC  
 CTCTGACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAG  
 TGTGGAGACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAA  
 ATGGGTTGGAACCCATCAGTGATCGCATATTTCATTGATGAGGGTTTGCTTGAGATAGAAAATG  
 GTGGCTCCTTTCTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTTCTTCTCAAGAGA  
 AAGTTGTAACCTCTCTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAAATAAAGAGTTCTTG  
 TTTCTGGGGGAA

**FIGURE 540**

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFADEKHPCKMAYFHELSSRVSFQEARLACES  
EGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSN  
SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAP  
VEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRKT  
SPNQSTLWISKSTRKESGMEV

**Important features:****Signal peptide:**

amino acids 1-21

**Transmembrane domain:**

amino acids 214-235

**N-glycosylation sites.**

amino acids 86-89 and 255-258

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 266-269

**N-myristoylation sites.**

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145  
and 212-217





**FIGURE 542**

MERAVRVESGVLVGVCVCLLLACPATATGPEVAQPEVDTTLGRVRGRQVGKGTDRLVNVFLGI  
 PFAQPPLGPDREFSAPHPAQPWEGVRDASTAPPMCLQDVESMNSSRFVLNGKQQIFSVSEDCLV  
 LNVYSPAIEVPAGSGRPVMVWVHGGALITGAATSYDGSALAAYGDVVVVTVQYRLGVLGFFSTG  
 DEHAPGNQGFLLDVVAALRWVQENIAPFGGDLNCVTVFGGSAGGSIIISGLVLSPPVAAGLFHRAI  
 TQSGVITTPGIIDSHPWPLAQKIANLTLACSSSSPAEMVQCLQQKEGEELVLSKKLKNTIYPLT  
 VDGTVPFKSPKELLKEKPFHSPVFLMGVNNHEFSWLIIPRGWGLLDTEQMSREDMLAISTPVL  
 TSLDVPPEMMPTVIDEYLGSSNSDAQAKCQAFQEFMGDVFINVPTVSFSRYLRDSGSPVFFYEF  
 QHRPSSFAKIKPAWVKADHGAEGAFVFGGPFLMDESSRLAFPEATEEEKQLSLTMMAQWTHFA  
 RTGDPNSKALPPWPQFNQAEQYLEINPVPRAGQKFREAWMQFWSETLPSKIQQWHQKQKNRKA  
 QEDL

**Important features:****Signal peptide:**

amino acids 1-27

**Transmembrane domain:**

amino acids 226-245

**N-glycosylation site.**

amino acids 105-109

**N-myristoylation sites.**

amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161,  
 162-168, 217-223, 227-233, 228-234, 232-238, 262-268, 357-363,  
 461-467

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 12-23

**Carboxylesterases type-B serine active site.**

amino acids 216-232

**FIGURE 543**

TGTGCGCTGGCCCTCGCCATG CAGACCCCGCGAGCGTCCCCTCCCCGCCCCGGCCCTCCTGCTTCTGCTGCTGCTA  
 CTGGGGGGCGCCACGGCCTCTTCTGAGGAGCCGCCCGCTTAGCGTGGCCCCCAGGGACTACCTGAACCAC  
 TATCCCGTGTGTTGTGGGCGAGCGGCGGCGGACGCTGACCCCGCAGAAGGTGCTGACGACCTCAACATCCAGCGA  
 GTCTGCGGGTCAACAGGACGCTGTTCATTGGGGACAGGGACAACCTCTACCGCGTAGAGCTGGAGCCCCCAGC  
 TCCACGGAGCTGCGGTACCAGAGGAAGCTGACCTGGAGATCTAACCCAGCGACATAAACGTGTGTGCGATGAAG  
 GGCAAACAGGAGGGCGAGTGTGAAACTTCGTAAAGGTGCTGCTCCTTCGGGACGAGTCCACGCTCTTTGTGTGC  
 GGTTCACACGCTTCAACCCGGTGTGCGCAACTACAGCATAGACACCCTGCAGCCCGTCGGAGACAACATCAGC  
 GGATATGGCCCGCTGCCCGTACGACCCCAAGCACGCCAATGTTGCCCTCTTCTCTGACGGGATGCTCTTCACAGCT  
 ACTGTTACCGACTTCTAGCCATTGATGCTGTCTATCTACCGCAGCCTCGGGGACAGGCCCCACCTGCGCACCGTG  
 AAACATGACTCCAAGTGGTTCAAAGAGCCTTACTTTGTCCATGCGGTGGAGTGGGGCAGCCATGTCTACTTCTTC  
 TTCCGGGAGATTGCGATGGAGTTTAACTACCTGGAGAAGGTGGTGGTGTCCCGCGTGGCCCCGAGTGTGCAAGAAC  
 GACGTGGGAGGCTCCCCCGCGTGTGGAGAAGCAGTGGACGTCTTCTGAAGGCGCGGCTCAACTGCTCTGTA  
 CCCGGAGACTCCCATTTCTACTTCAACGTGCTGCAGGCTGTACGGGCGTGGTCAGCCTCGGGGGCGGGCCCGTG  
 GTCCTGGCCGTTTTTTCCACGCCCAGCAACAGCATCCCTGGCTCGGCTGTCTGCGCCTTTGACCTGACACAGGTG  
 GCAGCTGTGTTTGAAGGCGCTTCCGAGAGCAGAAGTCCCCGAGTCCATCTGGACGCCGGTGCCGGAGGATCAG  
 GTGCCTCGACCCCGGCCCGGGTGTGCGCAGCCCCCGGGATGCAGTACAATGCCTCCAGCGCCTTGCCGGATGAC  
 ATCCTCAACTTTGTCAAGACCCACCCTCTGATGGACGAGGCGGTGCCCTCGCTGGGCCATGCGCCCTGGATCTG  
 CGGACCCTGATGAGGCACCAGCTGACTCGAGTGGCTGTGGACGTGGGAGCCGGCCCCCTGGGGCAACCAGACCGTT  
 GTCTTCTGGGTTCTGAGGCGGGGACGGTCTCAAGTTCCTCGTCCGGCCCAATGCCAGCACCTCAGGGACGTCT  
 GGGCTCAGTGTCTTCTGGAGGAGTTTGAAGCCTACCGGCCGGACAGGTGTGGACGGCCCCGGCGGTGGCGAGACA  
 GGGCAGCGGCTGCTGAGCGTCAGCCGCTGGGCGAGCGCCCGGGTATGCAGTACAATGCCTCCAGCGCCTTGCCGGATGAC  
 CGAGTGCCTGTGGCTCGCTGAGCGTCAGCCGCTGGGCGAGCGCCCGGGTGTGGCTGCCTTCCCCCGCTGCGTGGTC  
 GGGTGGGCCCCCGACGGCTCCTGCATCTTCTCAGCCCGGGCACCAGAGCCGCTTTGAGCAGGACGTGTCCGG  
 GCCAGCACCTCAGGCTTAGGGGACTGCACAGGACTCCTGCGGGCCAGCCTCTCCGAGGACCGCGCGGGGCTGGTG  
 TCGGTGAACCTGCTGGTAACGTGCTGCGTGGTGGCGGCCCTTCGTGGTGGGAGCCGTGGTGTCCGGCTTCAGCGTGGGC  
 TGTTTCGTGGGCTCCGTGAGCGGCGGAGCTGGCCCGGCGCAAGGACAAGGAGGCCATCCTGGCGCACGGGGCG  
 GCGAGGCGGTGCTGAGCGTCAGCCGCTGGGCGAGCGAGGCGCAGGCTCCCGGGGGCCGGGGCGGAGGCGGT  
 GCGGTGGCGCGGGGTTCCCCCGAGGCCCTGCTGGCGCCCTGATGCAGAACGGCTGGGCCAAGGCCACGCTG  
 CTGCAGGGCGGGCCCCACGACCTGGACTCGGGGCTGCTGCCACGCCCGAGCAGACGCCGCTGCCCGAGAAGCGC  
 CTGCCACTCCGCACCCCGACCCCGACGCCCTGGGCCCGCGCCTGGGACCACGGCCACCCCTGCTCCCGGCC  
 TCCGTTTATCTCTCTCTGCTGCTGGCGCCCGCCGGGCCCCCGAGCAGCCCCCGCGCCTGGGGAGCCGACC  
 CCGACGGCGCCTTATGCTGCCCGGCCCGGCCGCGCTCCACGGCGACTTCCCGCTACCCCCACGCCAGC  
 CCGGACCGCGCGGGTGGTGTCCGCGCCACGGGCCCTTGAGCCAGCCTCAGCCCGCGATGGCTCCCCGCG  
 CCTGGAGCCCCCCCCGACGGGCGAGCTGAGGAGGCCACTGGGCCCCACGCCCTCCGGCCGCCACCCTGCGC  
 CGACCCACACGTTCAACAGCGGCGAGGCCCGGCTGGGGACCGCCACCGCGGCTGCCACGCCCGGGCGGACA  
 GACTTGGCCCACCTCCTCCCCATGGGGGGCGGACAGGACTGCGCCCCCGTGGCCTAGGCGGGGGCCCCCG  
 ATGCCTTGGCAGTGCCAGCCACGGGAACCAGGAGCGAGAGCGGTGCCAGAACGCCGGGGCCCGGGGCAACTCCG  
 AGTGGGTGCTCAAGTCCCCCGGACCCACCGCGGAGTGGGGGGCCCCCTCCGCCACAAGGAAGCACAAACCAG  
 CTCGCCCTCCCCCTACCCGGGGCGCAGGACGCTGAGACGTTTGGGGTGGGTGGGCGGGAGGACTTTGCTATG  
 GATTTGAGGTTGACCTTATGCGCGTAGGTTTGGTTTGTGAGTTTGGTTTCTTTTGGGTTTCTTTGCGGTTTCTAACC  
 AATTGCACAACCTCGTTCTCGGGGTGGCGGCAGGAGGGGAGGCTTGGACGCCGTGGGGAAATGGGGGGCCACAG  
 CTGCAGACCTAAGCCCTCCCCACCCCTGGAAAGTCCCTCCCCAACCCAGGCCCTGGCGTGTGTGGGTGTGCG  
 TGCGTGTGCGTGGCGTTCGTGTGCAAGGGGCGGGGAGGTGGGCGTGTGTGCGTGCCAGCGAAGGCTGCTG  
 TGGCGTGTGTCAAGTGGGCCACGCTGCAGGTTGTGTGCCAGGCGACGATCGTGGTGGCCCCAGCGGCC  
 TGGGCGTTGGCTGAGCCGACGCTGGGGCTTCCAGAAGGCCCGGGGCTCTCCGAGGTGCCGGTTAGGAGTTTGAAC  
 CCCCCCACTCTGCAGAGGGAAGCGGGGACAATGCCGGGGTTTCAGGCAGGAGACACGAGGAGGGCCTGCCCGGA  
 AGTCACATCGGCAGCAGCTGTCTAAAGGCTTGGGGGCGGGGGCGGCAAGGTGGGTGGGGCCCCCTCGGTAA  
 ATACGGCCCCAGGGTGGTGAGAGAGTCCCATGCCACCCGTCCCCTTGTGACCTCCCCCTATGACCTCCAGCTGA  
 CCATGCATGCCACGTGGCTGGCTGGGTCTCTGCCCTCTTTGGAGTTTGCTCCCCCAGCCCCCTCCCCATCAAT  
 AAAACTCTGTTTACAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 544**

MQTPRASPPRPALLLLLLLLLLLGGAHGLFPEEPPPLSVAPRDYLNHYPVFVGSGPGRLTPAEGAD  
 DLNIQVRVLRVNRTLFIGDRDNLYRVELEPPTSTELRYQRKLTWRSNPSDINVCRMKGKQEGEC  
 RNFVKVLLLRDESTLFVCGSNAFNPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALFSDG  
 MLFTATVTDFLAIDAVIYRSLGDRPTLRITVKHDSKWFKEPYFVHAVEWGS HVYFFFREIAMEF  
 NYLEKVVVSRVARVCKNDVGGSPRVLEKQWTSFLKARLNCSVPGDSHFYFNVLQAVTG VVSLG  
 GRPVVLAVFSTPSNSIPGSAVCAFDLTQVA AVFEGRFREQKSPESIWTPVPEDQVPRPRPGCC  
 AAPGMQYNASSALPDDILNFVKTHPLMDEAVPSLG HAPWILRTLMRHQLTRVAVDVGAGPWGN  
 QTVVFLGSEAGTVLKFLVRPNASTSGTSGLSVFLEEFETYRPDRCGRPGGETGQRLLSLELD  
 AASGGLLA AFPRCVVRVPVARCQQYSGCMKNCIGSQDPYCGWAPDGSCIFLSPGTRAAFEQDV  
 SGASTSGLGDC TGLLRASLSEDRAGLVSVNLLVTSSVAAFVVGAVVSGFSVGFVGLRERREL  
 ARRKDKEAILAHGAGEAVLSVSRLGERRAQGP GGRGGGGGGGAGVPPEALLAPLMQNGWAKAT  
 LLQGGPHDLDSGLLPTPEQTPLPQKRLPTPHPHPHALGPRAWDHGHPLL PASASSSLLLLAPA  
 RAPEQPPAPGEPTPDGRLYAARPGRASHGDFPLTPHASPDRRRVVSAPTGPLDPASAADGLPR  
 WSPPPPTGSLRRPLGPHAPPAATLRRTHTFNSGEARPGDRHRGCHARPGTDLAHL LPYGGADR  
 TAPPVP

**Important features:****Signal peptide:**

amino acids 1-25

**Transmembrane domains:**

amino acids 318-339, 598-617

**N-glycosylation sites.**

amino acids 74-78, 155-159, 167-171, 291-295, 386-390, 441-445,  
 462-466

**Glycosaminoglycan attachment sites.**

amino acids 51-55, 573-577

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 102-106

**N-myristoylation sites.**

amino acids 21-27, 50-56, 189-195, 333-339, 382-388, 448-454,  
 490-496, 491-497, 508-514, 509-515, 531-537, 558-564, 569-575,  
 574-580, 580-586, 610-616, 643-649, 663-669, 666-672, 667-673,  
 668-674, 669-675, 670-676, 868-874, 879-885

**FIGURE 545**

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGG  
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAA  
ACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT  
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG  
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTTTCACGGGAG  
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAAGGCCTAAGATGAAAGCCTCTAGT  
CTTGCCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG  
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT  
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT  
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC  
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC  
AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCCACATGACA  
TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG  
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAG  
GAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT  
TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT  
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGATTGTCTTTATGCATCCCC  
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT  
TATTAGTTAATATATTTATTTATTTTTTTGCTATTTAATGTATTTATTTTTTTTACTTGGACATG  
AAACTTTAAAAAAATTACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTTAT  
ACAGTAAAAAATAACCTTGTAATTCTAGAAGAGTGGCTAGGGGGGTATTTCATTTGTAT  
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC  
TACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG  
ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG  
CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAAA





**FIGURE 548**

MFGTLLLYCFFLATVPALAETGGERQLSPEKSEIWGPGLKADVVLPAFYFYIQAVDTSGNKFT  
SSPGEKVFQVKVSAPEEQFTRVGVQVLDRKDGSFIVRYRMYASYKNLKVEIKFQGQHVAKSPY  
ILKGPVYHENCDCPLQDSAAWLREMNCPETIAQIQRDLAHFPAVDPEKIAVEIPKRFGQRQSL  
CHYTLKDNKVYIKTHGEHVGFRIFMDAILLSLTRKVKMPDVELFVNLGDWPLEKKKSNSNIHP  
IFSWCGSTDSKDIVMPTYDLTDSVLETMGRVSLDMMSVQANTGPPWESKNSTAVWRGRDSRKE  
RLELVKLSRKHPOLIDAAFTNFFFFKHDENLYGPIVKHISFFDFFKHKYQINIDGTVAAYRLP  
YLLVGDSVVLKQDSIYYEHFYNELQPWKHYIPVKSNLSDLLEKLKWAKDHDEEAKKIAKAGQE  
FARNNLMGDDIFCYFVKLFQEYANLQVSEPQIREGMKRVEPQTEDDLFPCTCHRRKTKDEL

**Important features:****Signal peptide:**

amino acids 1-17

**N-glycosylation sites.**

amino acids 302-306, 414-418

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 243-247, 495-499

**Tyrosine kinase phosphorylation site.**

amino acids 341-348

**N-myristoylation sites.**

amino acids 59-65, 118-124, 184-190, 258-264, 370-376, 439-445

**Endoplasmic reticulum targeting sequence.**

amino acids 499-504



**FIGURE 549**

GGGTGATTGAACTAAACCTTCGCCGCACCGAGTTTGCAGTACGGCCGTCACCCGCACCGCTGC  
CTGCTTGCGGTTGGAGAAATCAAGGCCCTACCGGGCCTCCGTAGTCACCTCTCTATAGTGGGC  
GTGGCCGAGGCCGGGGTGACCCTGCCGGAGCCTCCGCTGCCAGCGAC**ATG**TTCAAGGTAATTC  
AGAGGTCCGTGGGGCCAGCCAGCCTGAGCTTGCTCACCTTCAAAGTCTATGCAGCACCAAAAA  
AGGACTCACCTCCCAAAAATTCCGTGAAGGTTGATGAGCTTTCACTCTACTCAGTTCCCTGAGG  
GTCAATCGAAGTATGTGGAGGAGGCAAGGAGCCAGCTTGAAGAAAGCATCTCACAGCTCCGAC  
ACTATTGCGAGCCATACACAACCTGGTGTGAGGAAACGTACTCCCAAACTAAGCCCAAGATGC  
AAAGTTTGGTTCAATGGGGGTTAGACAGCTATGACTATCTCCAAAATGCACCTCCTGGATTTT  
TTCCGAGACTTGGTGTATTGGTTTTGCTGGCCTTATTGGACTCCTTTTGGCTAGAGGTTCAA  
AAATAAAGAAGCTAGTGTATCCGCCTGGTTTCATGGGATTAGCTGCCTCCCTCTATTATCCAC  
AACAAGCCATCGTGTTTGGCCAGGTCAGTGGGGAGAGATTATATGACTGGGGTTTACGAGGAT  
ATATAGTCATAGAAGATTTGTGGAAGGAGAACTTTCAAAGCCAGGAAATGTGAAGAATTCAC  
CTGGAAC**TAG**AAACTCCATGCTCTGCCATCTTAATCAGTTATAGGTAAACATTGGAAA  
CTCCATAGAATAAATCAGTATTTCTACAGAAAAATGGCATAGAAGTCAGTATTGAATGTATTA  
AATTGGCTTTCTTCTTCAGGAAAACTAGACCAGACCTCTGTTATCTTCTGTGAAATCATCCT  
ACAAGCAAACCTGGAATCCCTTCACCTAGAGATAATGTACAAGCCTTAGAACTCCTCAT  
TCTCATGTTGCTATTTATGTACCTAATTAAAACCCAAGTTTAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAA

## **FIGURE 550**

MFKVIQRSVGPASLSLLTFKVYAAPKKDSPPKNSVKVDELSLYSVPEGQSKYVEEARSQLEES  
ISQLRHYCEPYTTWCQETYSQTKPKMQSLVQWGLDSYDYLQNAPPGFFPRLGVIGFAGLIGLL  
LARGSKIKKLVYPPGFMGLAASLYYPQQAIVFAQVSGERLYDWGLRGYIVIEDLWKENFQKPG  
NVKNSPGTK

### **Important features:**

#### **Signal peptide:**

Amino acids 1-23

#### **Transmembrane domain:**

Amino acids 111-130

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 26-30

#### **Tyrosine kinase phosphorylation site:**

Amino acids 36-44

#### **N-myristoylation sites:**

Amino acids 124-130;144-150;189-195